

“十二五”普通高等教育本科国家级规划教材

普 通 高 等 教 育 精 品 教 材

配套用书



21世纪大学本科计算机专业系列教材

吴功宜 吴英 编著

计算机网络教师用书（第4版）

http://www.tup.com.cn

- 根据教育部“高等学校计算机科学与技术专业规范”组织编写
- 与美国 ACM 和 IEEE CS *Computing Curricula* 最新进展同步
- 国家级精品教材配套用书
- 全国高校出版社优秀畅销书配套用书

清华大学出版社

“十二五”普通高等教育本科国家级规划教材
21世纪大学本科计算机专业系列教材

计算机网络教师用书(第4版)

吴功宜　吴　英　编著

清华大学出版社
北　京

内 容 简 介

《计算机网络教师用书(第4版)》对主教材的知识体系、每章的知识点结构,以及内容前后衔接关系均做出了分析,以帮助任课教师准确把握全局与局部内容。作者总结了多年教学科研工作中遇到的问题,按照主教材的章节顺序,站在初学者的角度,提出了269个“为什么”,并逐一做了回答;还对一些重要和容易混淆的技术术语做了分析与比较。教师用书对主教材每章较难的练习题给出了答案,供任课教师参考。同时,也为任课教师在备课过程中会遇到的问题,以及拓展知识面提供帮助。

本书可以作为计算机、软件工程、信息安全、物联网工程、通信工程与电子信息等专业的计算机网络、数据通信技术及相关课程的教师参考书,也可以作为计算机、电子信息类专业的本科生、研究生与工程技术人员的参考书。

图书在版编目(CIP)数据

计算机网络教师用书/吴功宜,吴英编著.—4版.—北京:清华大学出版社,2017
(21世纪大学本科计算机专业系列教材)
ISBN 978-7-302-46901-8

Ⅰ.①计…　Ⅱ.①吴…　②吴…　Ⅲ.①计算机网络－高等学校－教学参考资料　Ⅳ.①TP393

中国版本图书馆CIP数据核字(2017)第053319号

责任编辑:张瑞庆　薛　阳
封面设计:何凤霞
责任校对:梁　毅
责任印制:李红英

出版发行:清华大学出版社
网　　址:http://www.tup.com.cn, http://www.wqbook.com
地　　址:北京清华大学学研大厦A座　　**邮　　编**:100084
社 总 机:010-62770175　　**邮　　购**:010-62786544
投稿与读者服务:010-62776969, c-service@tup.tsinghua.edu.cn
质量反馈:010-62772015, zhiliang@tup.tsinghua.edu.cn
课件下载:http://www.tup.com.cn,010-62795954
印 装 者:清华大学印刷厂
经　　销:全国新华书店
开　　本:185mm×260mm　　**印　　张**:23.75　　**字　　数**:574千字
版　　次:2004年8月第1版　2017年6月第4版　　**印　　次**:2017年6月第1次印刷
印　　数:1～1000
定　　价:59.00元

产品编号:069470-01

本书主审：钱德沛

再版前言

FOREWORD

作者在40年前大学毕业留校工作时，曾经向一位老先生请教如何教好一门课。老先生送给作者的一句话是："你要给学生一勺水，那么你自己就需要准备一桶水"。多年的教学实践让作者深刻地体会到：要教好一门课，教师需要通过教学研究，深入理解课程的精髓，掌握教学规律；通过科研来提高自身的学术水平，提高理论联系实际的能力。此外，还需要在教学的过程中，不断地向学生学习，了解初学者对某些问题的认识过程与规律，搞懂初学者对哪些问题不容易掌握，以及为什么不容易掌握。

作者在规划教材体系建设时，注意到理论知识学习和实际能力培养的关系问题。主教材内容定位与网络技术发展的总体水平，力求做到知识体系设计合理，难易程度适中，配套教材齐全，能够适应我国不同地区、不同学校和专业网络课程教学的基本要求。经过十多年的努力，基本形成了由"一本主教材、四本辅助教材和一个电子教案"构成的教材体系，为计算机网络课程教学改革提供了一个良好的平台。

主教材《计算机网络(第4版)》具有以下特点。

第一，打牢网络理论与技术基础。

在知识结构的设计中，坚持以支撑互联网、移动互联网与物联网发展的共性技术为主线，每一章内容力求集中回答其中一个基本的问题。

第1章　计算机网络概论：什么是计算机网络?

第2章　物理层：网络中比特流的传输是如何实现的?

第3章　数据链路层：网络中数据传输的正确性是如何保证的?

第4章　介质访问控制子层：最常用的 Ethernet 与 Wi-Fi 的网络功能是如何实现的?

第5章　网络层：网络互联是如何实现的?

第6章　传输层：网络环境中分布式进程通信是如何实现的?

第7章　应用层：网络应用系统是如何设计和实现的?

第8章　网络安全：如何保证网络安全?

通过剖析常用的网络应用实例，对网络应用系统的设计与实现方法进行讨论，帮助读者渐进、潜移默化地接受前人成熟的研究方法与成果，为继续学习和研究网络技术奠定基础。

第二，贴近技术发展前沿。

计算机网络是当今计算机科学与技术学科中发展最为迅速的技术之一，也是计算机应用中一个空前活跃的领域。如果说广域网的作用是扩大了信息社会中资源共享的广度，城域网是扩大了用户接入互联网的范围，局域网是扩大了信息资源共享的深度，个人区域网络

是增强了人类共享信息的灵活性,那么物联网是在互联网技术的基础上,利用RFID、感知技术与无线传感器网络自动获取物理世界的信息,构建“人机物”深度融合的系统。在物联网时代,计算机、智能手机、传感器、射频标签、机器人、可穿戴设备都会连接到网络之中。计算机网络技术的发展为物联网“人机物”深度融合系统的形成提供了技术支撑;而物联网技术与产业的发展又为计算机网络技术提出了很多富有挑战性的研究课题。作者潜心研读了近年来国内外计算机网络的重要著作、文献,并结合个人与科研团队的研究工作,贴近技术发展前沿,探讨计算机网络知识体系中“变”与“不变”的关系,分析学生学习过程中的“难点”和为什么会成为难点的问题,突出计算机专业的特点,对相关章节内容做出了适当地取舍,相对较多地增加了无线网络的分量,形成第4版的知识体系。

第三,以“系统观”的思路组织网络知识体系。

计算机专业学生需要更强调计算机系统能力的培养。计算机专业学生系统能力的核心是培养学生具有设计和构建以计算技术为核心、新的应用系统的能力,而网络知识是计算机系统能力的重要组成部分。因此,计算机专业学生系统能力的培养要将计算机与计算机网络看成一个有机的整体,引导学生用计算机组成原理、操作系统的基础知识,去理解计算机网络基本工作原理,学会用软件编程的方法去实现网络服务功能,使得学生能够准确描述与构建出真实网络应用系统的模型,以及有效地构造网络应用系统。教程在组织每一章、每一个知识点,以及习题、网络软件编程与硬件训练时,都力求做到这一点。

“应用层”选取了有代表性的Internet应用(FTP应用),从体系结构、网络协议、软件编程与操作系统进程通信的角度,采用“系统观”的方法对计算机网络原理与实现方法,从硬件与软件实现的角度进行了概括和总结。

第四,贯彻“以能力培养为导向”的教学理念。

计算机网络是一门应用性与实践性很强的课程。学生只有通过系统地训练,才有可能真正掌握和深入理解网络技术的基本理论与方法。教学团队在规划教材体系建设时,坚持“以能力培养为导向”的指导思想,经过近二十年的努力,基本形成了由“一本主教材、四本辅助教材和一个电子教案”构成的教材体系。与主教材配套的有《计算机网络教师用书(第4版)》《计算机网络实验指导书(第3版)》及《计算机网络软件编程指导书》。

作为配套教材之一的《计算机网络教师用书(第4版)》仍然保持着以下几个主要特点。

第一,教师用书对整个课程的知识体系,每章的知识点结构、内容的前后衔接均做出了分析,帮助任课教师准确地把握全局与局部内容的关系。

第二,作者总结了多年来在计算机网络教学、科研工作中遇到的问题,学生提出的问题,以及同行之间交流的问题,针对主教材各章节重要的知识点、难点,按照章节顺序,提出了269个“为什么”,并逐一做了回答;通过对重要和容易混淆技术术语的“辨析”,展开分析和比较,希望引起更深入的讨论。书中提出的269个问题需要有一个日积月累的过程。把这些问题积累起来,再整理出来,确实经历了漫长和多次反复的过程。

第三,对教材中做了“减”法处理,简化掉一些过渡性、过时的技术,教师在备课中全面地了解技术的演变过程,教师用书拿出一定的篇幅,对这部分知识做了一些讨论。主教材写得比较清晰和比较易懂的内容,教师用书就尽量简化。教师用书中的很多内容的选择是希望能够帮助任课教师加深对某些概念的理解,丰富相关的背景知识。每位老师的知识结构与教学经验不同,教学对象也不相同,因此建议各位老师有选择地阅读教师用书的相关内容。

第四，为了帮助老师组织好理论教学，教师用书给出了主教材每章较难的练习题的答案，供任课教师参考。习题解析一书将对相关的习题做出较为详细的解析。

作者认为，教师一辈子最主要的任务是尽自己最大的努力，把承担的每门课程教好，不能误人子弟。作为一名好的教师，他在教授这门课程时最重要的是“知道学生在哪些知识的学习时，可能有哪些知识点不容易掌握，以及为什么不容易掌握”。教师最大的贡献是教出一批好学生。要达到这个目的，教师需要对自己所教的课程开展教学研究。作者选择了国际流行的教材 *Computer Networks*（Andrew S. Tanenbaum），结合计算机网络与 Internet 技术发展背景，对第 1 版到第 5 版的各个版本的特点进行了分析。作者重点比较和分析了第 3 版到第 5 版每一章节的具体内容，对每章内容的变化，以及变化的原因进行了解释。通过对 5 个版本教材内容的比较与分析，作者研究了计算机网络教学体系中“变”与“不变”的关系，并结合对技术与教学内容的理解，提出了主教材的整体结构。作者希望将这些教学研究成果奉献给同行，同时引起讨论；并希望从讨论中获取大家更加宝贵的教学经验，共同提高我国计算机网络课程的教学水平。从这个角度来看，《计算机网络教师用书（第 4 版）》只能起到“抛砖引玉”的作用。

本书的第一部分与第二部分的 0～6 章由吴功宜编写，第二部分的 7、8 章由吴英编写。1～9 章的习题答案由吴英提供，全书由吴功宜统编。

本书在写作过程中得到南开大学徐敬东、张建忠以及清华大学出版社的帮助与指导，在此表示感谢。同时要感谢这些年与作者交流，并提出很多问题的老师和同学们。由于作者水平有限，书中疏漏和不妥在所难免，敬请读者批评和指正。

吴功宜　wgy@nankai.edu.cn
吴　英　wuying@nankai.edu.cn
南开大学
计算机与控制工程学院
计算机与信息安全系
2016 年 10 月 26 日

目　录

CONTENTS

第 0 章 计算机网络课程知识点组织的总体说明

0.1 从科研工作角度探讨计算机网络课程改革的定位

0.1.1 我国互联网规模与互联网普及率的增长

计算机网络是当今计算机科学与技术学科中发展最为迅速的技术之一，也是计算机应用中一个空前活跃的领域。计算机网络技术对新一代信息技术与战略性新兴产业发展将会产生重要的推动作用。

从 1995 年开始的“九五”至“十二五”等 4 个五年计划、20 年的建设，我国社会信息化建设已经取得了重大的进展。截止到 2015 年年底，我国的互联网网民规模已经达到 6.88 亿，普及率达到 50.3%。图 0-1 给出了从 2000—2015 年我国互联网网民数量与互联网普及率增长的数据。

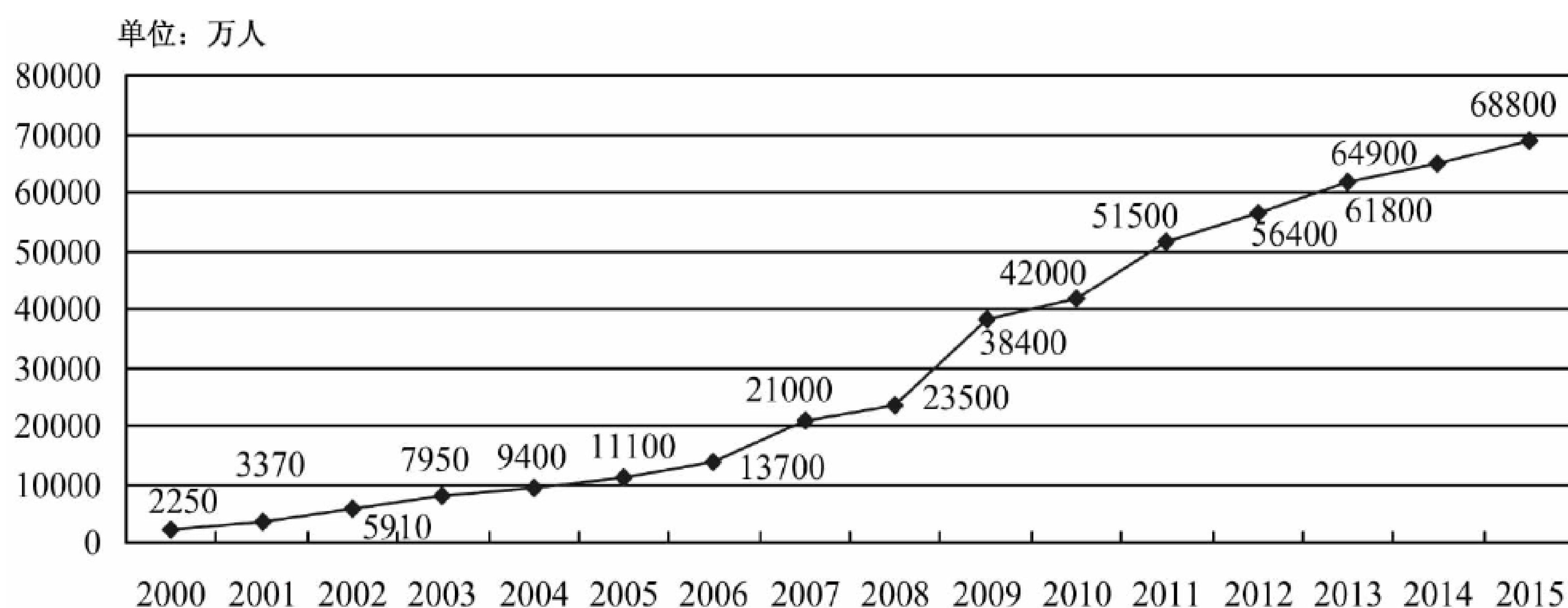
图 0-2 给出了从 2006 年至 2015 年我国手机网民数量和占互联网用户比例增长的数据。截止到 2015 年年底，我国的手机网民规模达到 6.2 亿，占网民总数的 90.1%。

随着我国经济的快速发展，我国 Internet 应用规模与水平将会以更快的速度发展，我国社会的经济、政治、文化、科技与教育对计算机网络与 Internet 技术水平要求将会越来越高，计算机网络与 Internet 应用发展的空间还是很大的，这也给计算机网络课程教学的改革提出了更高的要求。

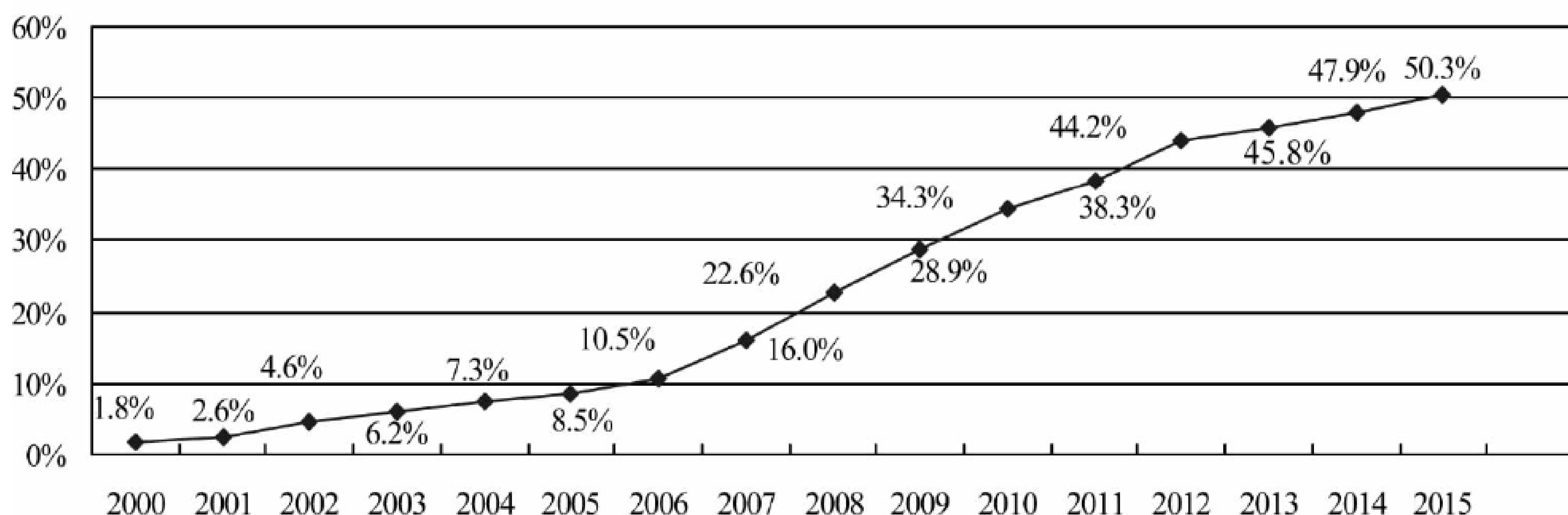
综上所述，我们可以清晰地得出的结论是：计算机网络正在沿着“互联网—移动互联网—物联网”的轨迹，“由小到大”地发展、壮大；“由表及里”地渗透到社会的各个角落；遵循“互联网＋”的模式，在与各行各业的跨界融合中，推动着我国国民经济发展方式的转型与社会发展。网络在社会经济、科技、文化进步中的作用越来越重要，网络知识已经成为计算机与信息技术相关专业大学生主要的知识基础之一，为国家培养网络人才的计算机网络课程也必须与时俱进地改革。

0.1.2 计算机网络从互联网、移动互联网到物联网的发展趋势分析

从事计算机网络研究的技术人员能够清晰地认识到：推动计算机网络技术与产业发展



(a) 我国互联网网民规模的增长



(b) 我国互联网普及率的增长

图 0-1　我国的互联网规模的增长

的动力来自两个方面，一是计算机产业；二是电信产业。在电信产业中最有影响力的国际组织是国际电信联盟(International Telecommunications Union，ITU)。20 世纪 90 年代，当互联网技术快速发展时，ITU 的高层研究人员已经前瞻性地认识到：互联网技术的广泛应用必将深刻地影响电信业的发展。ITU 的研究人员将互联网应用对电信业发展影响作为一个重要的课题开展研究，并从 1997 年至 2005 年发表了七份“ITU Internet Reports”系列研究报告(如图 0-3 所示)。

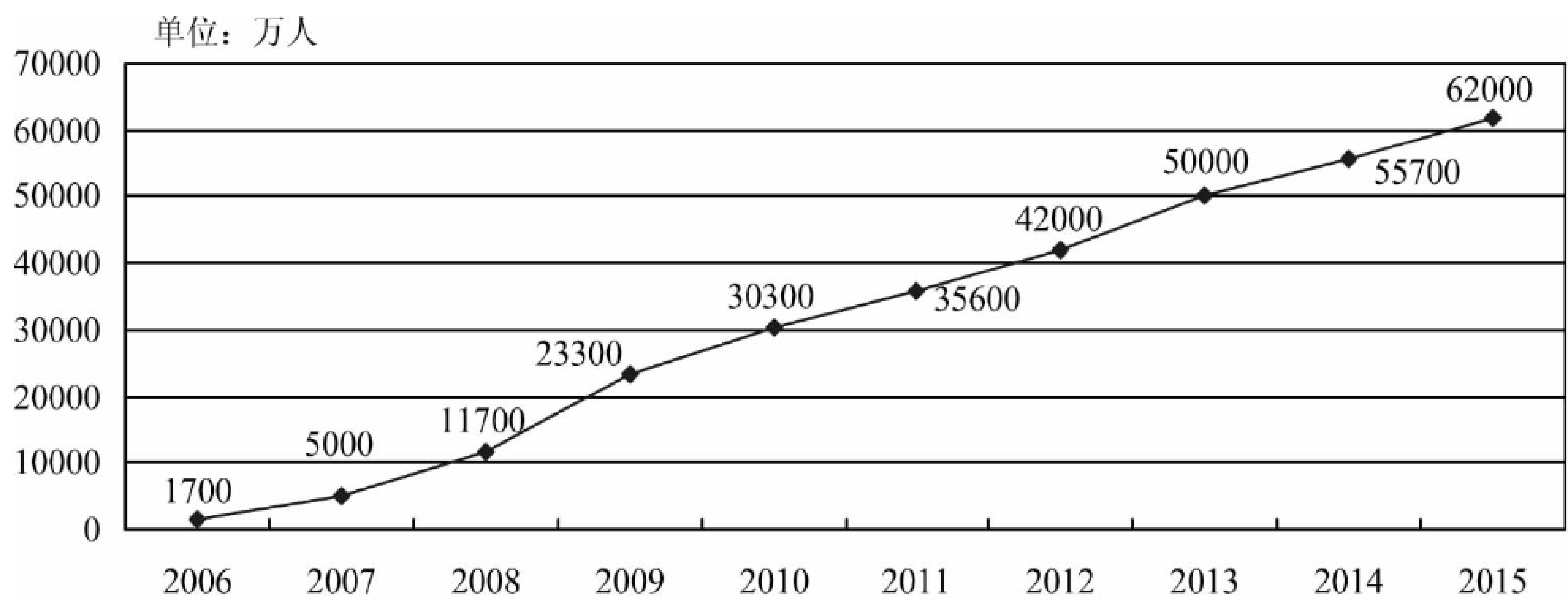
我们从这七份研究报告的内容中可以看出 ITU 对互联网的发展对国际电信业发展影响的判断，以及物联网概念提出的背景。

1. 1997 年：《挑战网络：电信和互联网》

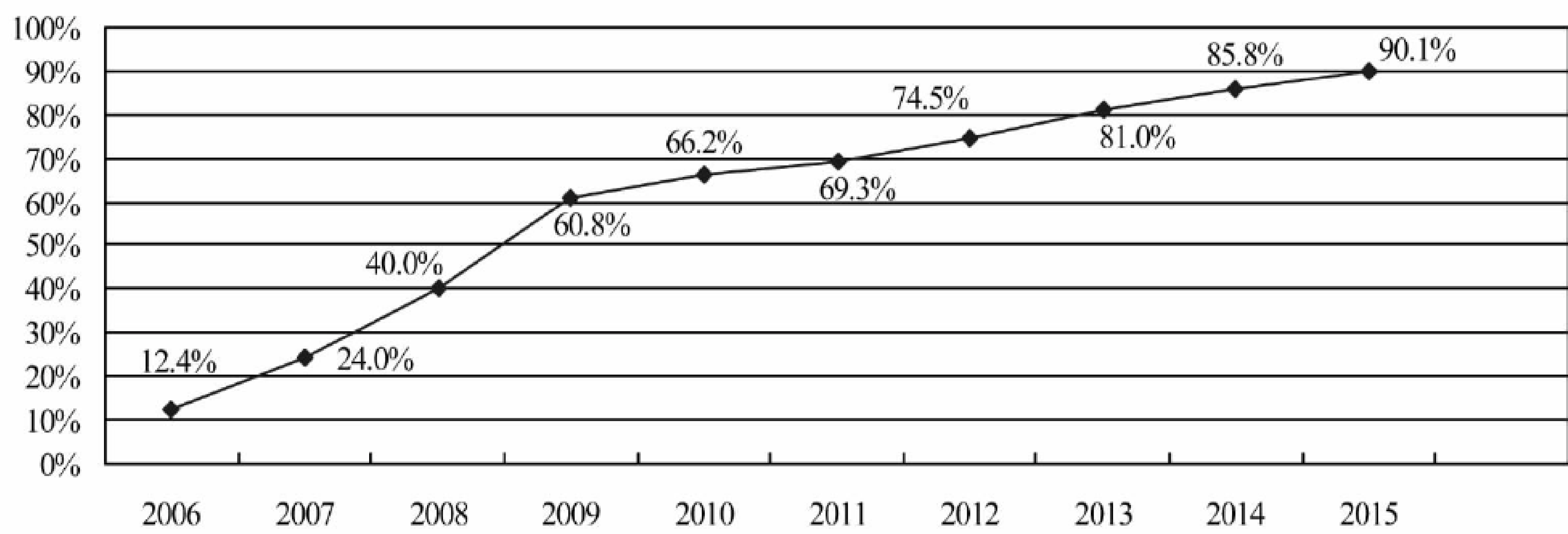
1997 年 ITU 发布的第 1 个研究报告的题目是《挑战网络：电信和互联网(*Challenges to the network*：*Telecoms and the Internet*)》。这份报告是为 1997 年 9 月 ITU 在日内瓦举行的电信展示与论坛会议(ITU TELECOM WORLD)准备的。报告论述了互联网的发展对电信业的挑战，同时指出互联网给电信业带来了重大的发展机遇。

2. 1999 年：《互联网发展》

1999 年 ITU 发布的第 2 个研究报告的题目是《互联网发展(*Internet for Development*)》。报告描述了互联网应用对于未来社会发展的影响，展望了互联网对促进人与人之间交流的作用，并就如何利用互联网帮助发展中国家发展通信事业进行了讨论。



(a) 我国手机网民数量的增长



(b) 手机网民数量占互联网用户比例的增长

图 0-2　我国的手机互联网规模的发展

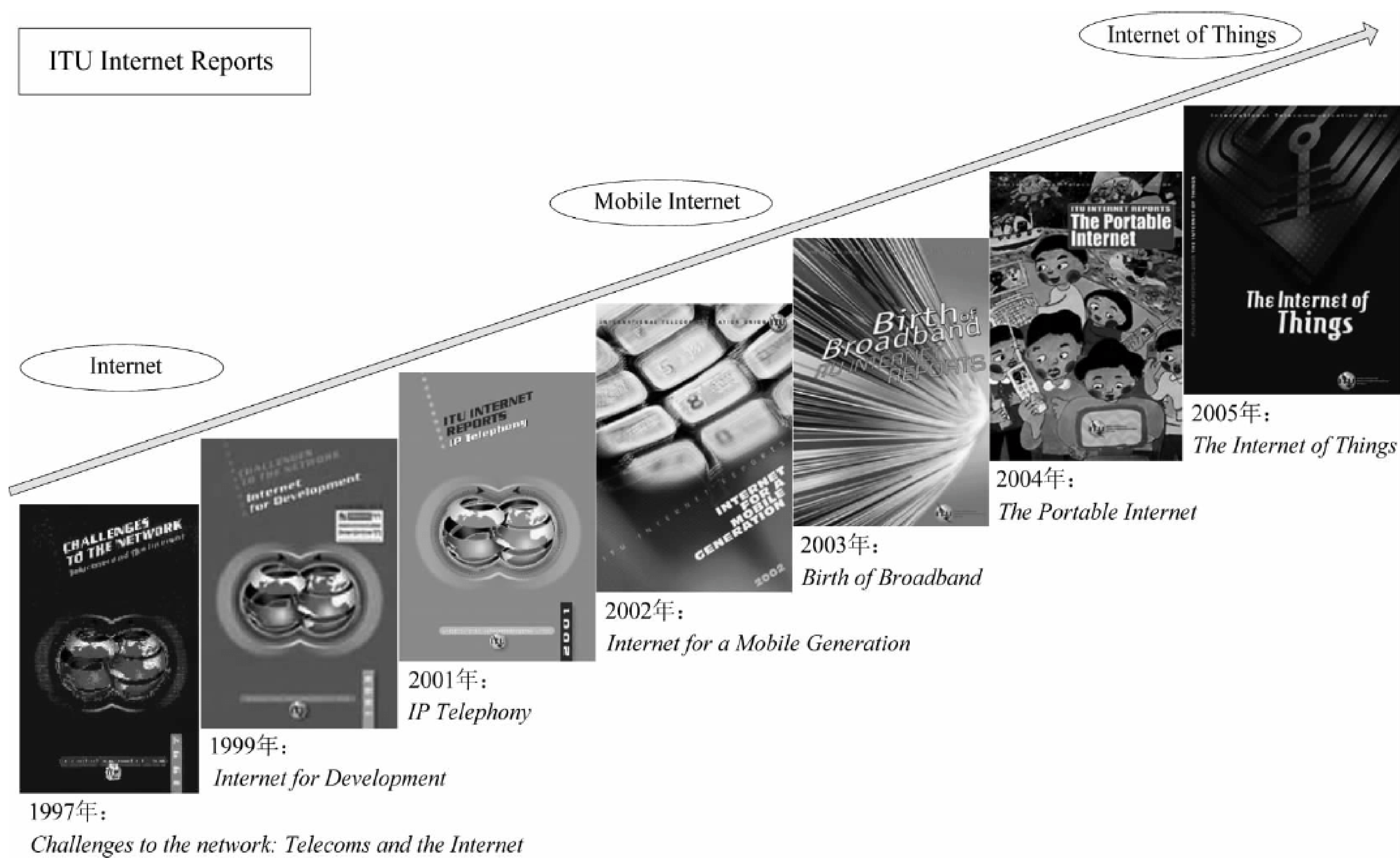


图 0-3　ITU 提出物联网概念的过程

3. 2001年:《IP电话》

2001年ITU发布的第3个研究报告的题目是《IP电话(*IP Telephony*)》。报告对IP电话技术标准、服务质量、带宽、编码与网络结构,并对IP电话应用领域、对电信运营商的传统电话业务的影响,并对IP电话监管问题进行了系统的讨论。

4. 2002年:《移动互联网时代》

2001年ITU发布的第4个研究报告的题目是《移动互联网时代(*Internet for a Mobile Generation*)》。报告讨论了移动互联网发展的背景、技术与市场需求,以及手机上网与移动互联网服务。报告指出:移动通信同互联网的融合,将构筑移动互联网美好的未来。移动互联网的发展将带领我们进入一个移动的信息社会。

5. 2003年:《宽带的诞生》

2003年ITU发布的第5个研究报告的题目是《宽带的诞生(*Birth of Broadband*)》。这份报告是专门为ITU于2003年10月在日内瓦举办的2003年世界电信展示会和论坛准备的。作为2003年电信产业的"热点"之一,宽带成为2003年展示会上的一大亮点。报告系统地介绍了宽带技术发展的过程,以及宽带技术对全世界电信业发展的影响。报告介绍了宽带网络发展比较好国家的成功案例,描述了宽带技术对未来信息社会的影响。同时,报告也讨论了计算机、通信和广播电视网络的三网融合问题,以及未来宽带网络发展动向,以及新的应用问题。

6. 2004年:《便携式互联网》

2004年ITU发布的第6个研究报告的题目是《便携式互联网(*The Portable Internet*)》。这份报告是专门为ITU于2004年9月11日在韩国釜山展开的2004年ITU亚洲电信展和论坛准备的。报告系统地讨论了应用于移动互联网的高速无线上网便携式设备的市场潜力、商业模式、发展战略与市场监管,以及移动互联网技术、市场的发展趋势,未来移动互联网技术的发展对信息社会的影响等问题。

7. 2005年:《物联网》

2005年ITU在突尼斯举行的"信息社会峰会"上发布了第7个研究报告——《物联网(*The Internet of Things*)》。报告描述了世界上的万事万物,小到钥匙、手表、手机,大到汽车、楼房,只要嵌入一个微型的RFID芯片或传感器芯片,通过互联网就能够实现物与物之间的信息交互,从而形成一个无所不在的"物联网"。世界上所有的人和物在任何时间、任何地点,都可以方便地实现人与人、人与物、物与物之间的信息交互。报告预见:RFID、传感器技术、智能嵌入式技术及纳米技术将广泛应用。

从这七份研究报告讨论的主题与内容中,我们可以清晰地从ITU专家的研究报告中看出:计算机网络经历了从互联网、移动互联网到物联网的发展与演变的过程。因此,依据国际权威机构的研究报告与产业发展趋势,我们将计算机网络发展过程归纳为:计算机网络形成、互联网、移动互联网与物联网等四个阶段是恰当的。

0.2 从互联网、移动互联网到物联网中网络技术的"变"与"不变"

面对快速发展的技术,我们只能从网络最基本的原理出发,总结提炼,让读者能够循序渐进地了解技术的发展过程,理解网络应用系统设计的基本思想。要做到这一点就需要处

理好网络课程中“变”与“不变”的基本关系问题。解决这个问题需要有科研与教学研究经验与成果作为支撑，需要做很多艰苦细致的分析、研究与取舍的工作。我们需要根据多年科学研究和教学工作的积累，研究从“互联网—移动互联网—物联网”的发展过程中，基本理论与核心技术的“变”与“不变”的关系，确定知识点结构与课程内容的取舍。

0.2.1 从计算机体系结构的角度去认识网络技术中的“变”与“不变”

从计算机体系结构的角度来看：计算机网络中连接计算机与网络的关键设备是网卡。对于主机来说，网卡与键盘、显示器、磁盘一样，都属于一种外部设备。在物联网中，RFID读写器、传感器与键盘、显示器、磁盘一样，也都属于主机的一种外部设备。图0-4给出了基于RFID物联网应用系统与计算机系统结构关系示意图。从图中可以看出，基于RFID应用系统的物联网结构只是在计算机网络的基础上，增加了RFID标签、RFID读写器，以及将读写器接入计算机系统的通信适配器，其他部分并没有出现实质性的变化，变化的部分主要表现在应用层。这个结论也可以推广到基于无线传感器的物联网系统，以及采用其他传感技术的物联网应用系统中。因此，从计算机体系结构的角度来看，互联网与物联网没有本质的区别。

0.2.2 从计算机操作系统的角度去认识网络技术中的“变”与“不变”

从网络层次结构与操作系统、硬件结构相关性的角度看，这里存在着两个重要的边界（如图0-4所示）。

（1）操作系统边界。

不同的操作系统在实现TCP/IP协议的方式可能有所不同，典型的结构如图0-5所示。从网络层次结构的角度，应用层与传输层、网络层软件之间“操作系统内部系统软件/操作系统之外应用软件”的界线。其中，实现传输层、网络层协议的软件属于操作系统的系统软件的一部分。而应用层实现应用层协议，提供各种网络服务功能的应用软件不属于操作系统的一部分。当用户启动某种网络应用，如Web浏览服务时，Web浏览器程序作为一个打开的进程，在本机操作系统的控制下，发起与Web服务器端的进程通信，执行HTTP协议规范，实现浏览Web页的功能。从传输层角度，网络的每一项服务都是对应一个“服务程序”进程。进程通信的实质是实现进程之间的相互作用。网络环境中的进程通信要解决的一个重要问题是确定进程间的相互作用模式。物联网应用系统的软件编程同样大量采用基于Web协议的B/S结构。

（2）协议地址边界。

从网络层次结构的角度，应用层采用的是域名、传输层采用的是TCP或UDP的端口号、网络层采用的是IP地址；数据链路层采用的是MAC硬件地址。因此，可以看出从“协议地址边界”的角度：网络层及以上采用的是软件地址，网络层以下使用的是硬件地址。

网络层与数据链路层之间存在着“IP地址/MAC地址”的协议地址边界。网络层一端使用通过软件设置的IP地址，而数据链路层使用的是全球唯一与固定不变的MAC地址，因此MAC地址也叫作“物理地址”。

从计算机操作系统与应用系统软件编程的角度看，互联网与物联网没有本质的区别。

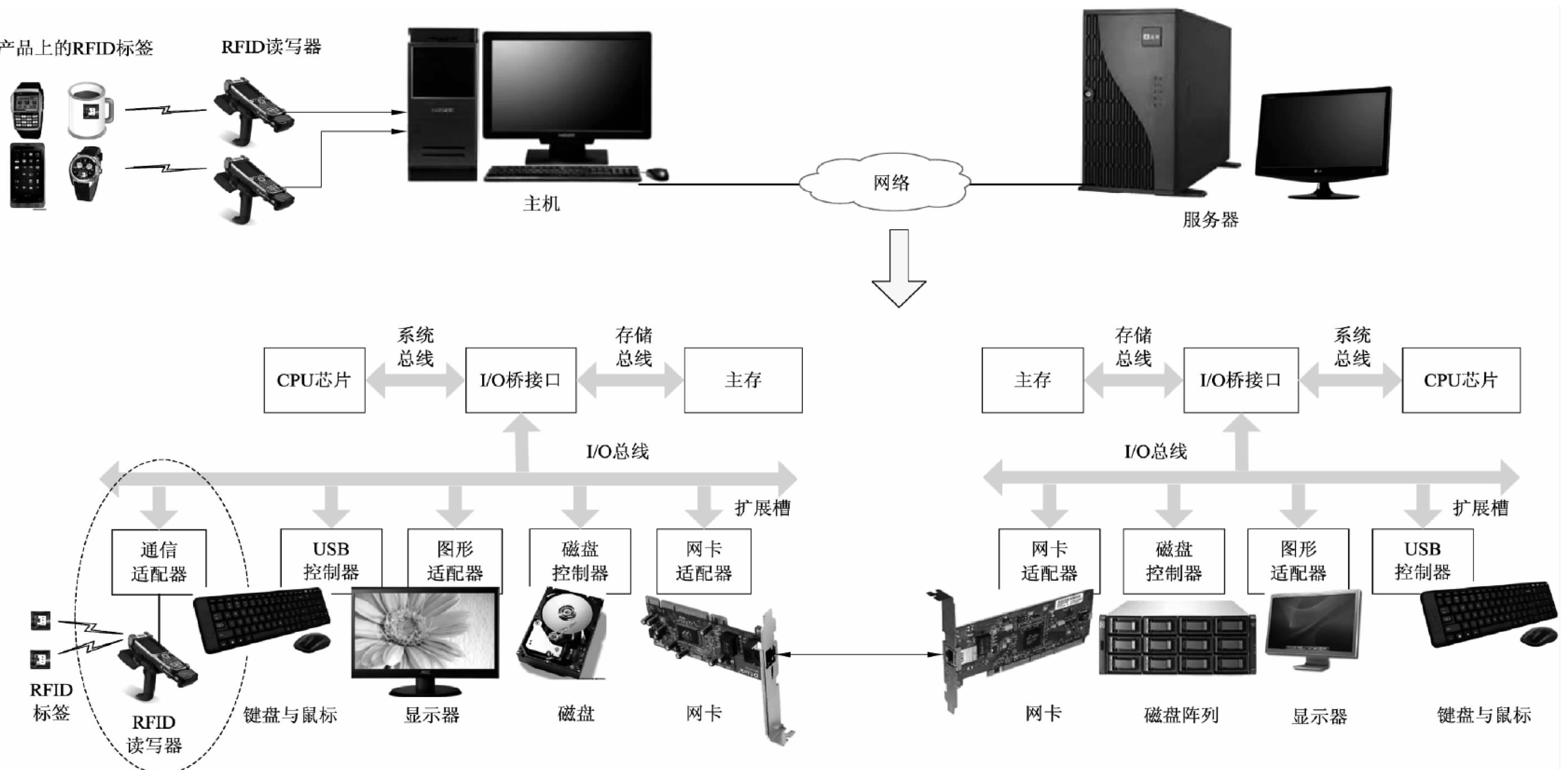


图 0-4　基于 RFID 物联网应用系统与计算机系统结构关系示意图

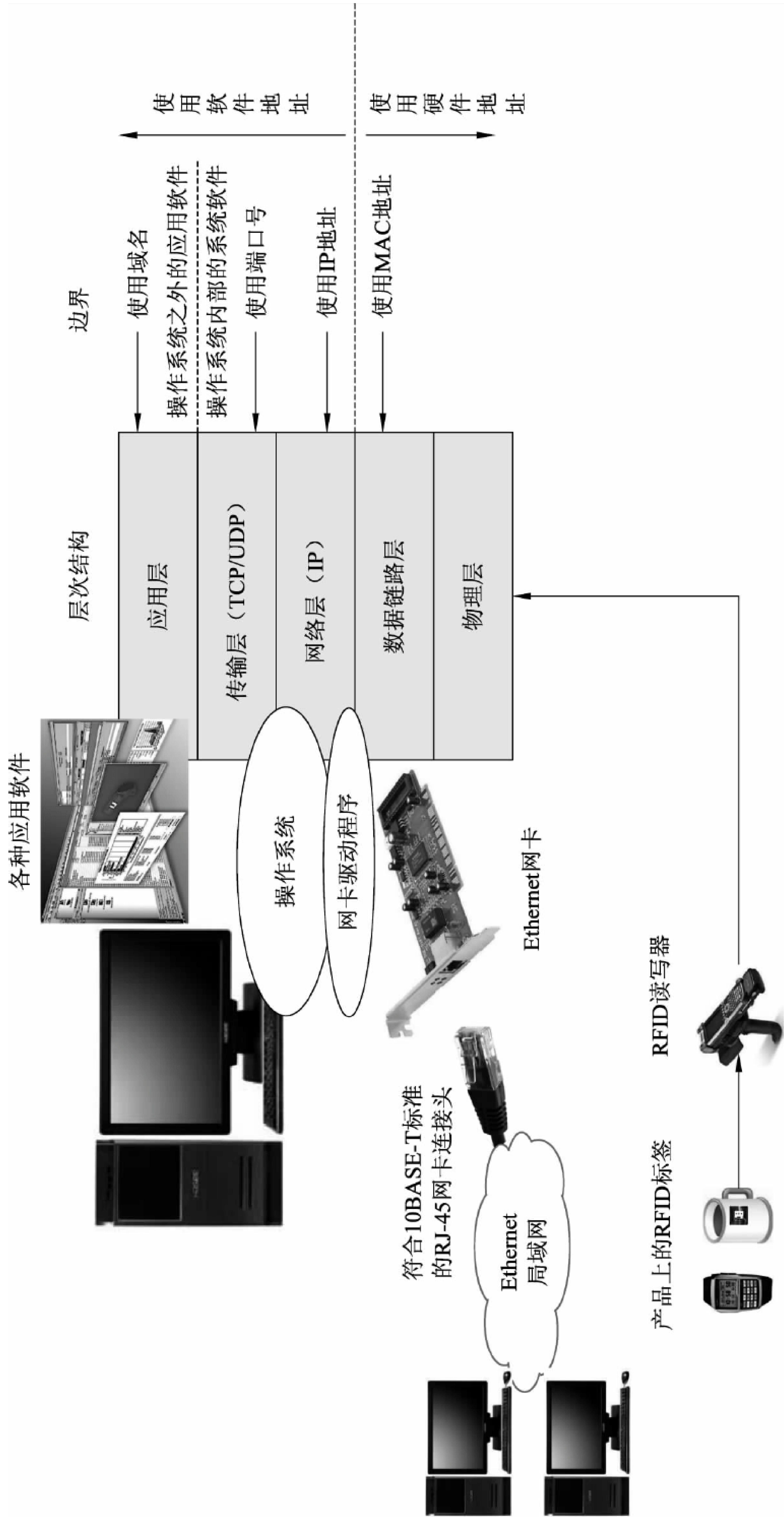


图 0-5 基于 RFID 物联网应用系统与操作系统关系示意图

0.2.3 从网络系统结构设计的角度去认识网络技术中的“变”与“不变”

我们可以通过一个覆盖全国的大型零售企业物联网应用系统为例,总结和归纳出支撑物联网应用系统的网络系统设计方法。

一个覆盖全国的大型零售企业是由总公司,分布在不同地区的分公司、仓库、配送中心,以及分散在不同位置的销售商店或超市组成。构建一个大型零售企业物联网应用系统一定需要将总公司局域计算机网络作为核心网络,通过核心路由器、光纤与汇聚路由器连接,实现与不同地区的分公司、仓库、配送中心的局域计算机网络的互联;汇聚路由器再通过光纤或其他通信线路,使用接入路由器,接入基层销售商店或超市的局域网,以形成覆盖全国的大型零售企业专用网络(如图0-6所示)。

大型零售企业专用网络是企业投资建设的,专门用于企业内部涉及商业机密的业务数据传输,因此外部互联网用户是不能够直接访问企业内部网络的。但是,大型零售企业也需要提供网上供应商的采购,以及针对网上用户的网购服务。因此,在公司网络系统的设计中,我们还可以设计一个公司外网。

公司外网需要通过防火墙等网络安全设备接入互联网,在防火墙内部设置Web、FTP与E-mail服务器,完成公司对外宣传、发布商品信息,向供应商采购货物的功能,同时要接受互联网客户的订单、投诉与售后服务功能。

从网络安全的角度出发,客户网上订购必须经过公司内部人员处理之后,通过连接公司外网与公司内网的代理服务器,由公司内部工作人员将客户订购信息转发到公司内网中专门用于处理网上定购的部门。

分散在不同城市、不同位置的销售商店或超市承担着商品销售的任务。按照物联网思路设计的销售商店或超市,它的商品贴有RFID标签。商店内部有固定的商品信息查询终端、商品信息显示屏,以及移动的导购机器人。购物的客户可以推着智能购物车,将她需要购买的商品放到智能购物车中,当她准备结账时,只要将智能购物车经过智能收款机时,只能收款机就已经通过商品的RFID标签告知客户一共需要付款的金额,发送到客户智能手机上。客户只要在通过RFID标签进行身份认证的智能手机上按一个确认键,整个购物的过程就非常方便地完成了。

贴在商品上的RFID标签、RFID读写器、嵌入RFID标签的智能手机,以及用于保障商店安全的视频传感器——摄像头,都属于具有自动感知能力的嵌入式电子设备;商品信息查询终端、商品信息显示屏、导购机器人都属于典型的嵌入式系统。因此,设计、实现与运行一个大型零售企业物联网应用系统的技术人员,必须懂得感知技术与嵌入式技术。

分散在不同位置的销售商店或超市需要将实时的销售数据传送到总公司网络存储与处理。总公司的工作人员使用数据挖掘技术,对实时数据进行智能分析,从中找出不同地区、不同商品的畅销、滞销的规律,对库存、缺货数据进行分析,制定不同地区的商品宣传、促销策略。

同时,分析人员需要根据不同地区不同商店的商品库存、缺货数据,货物配送中心的商品库存数据,以及货物配送车辆运送的货物品种、数量、当前位置,规划最佳配送方案,并将货物配送车辆最佳配送方案发送到不同的配送中心,由配送中心调度货物配送车辆的行车

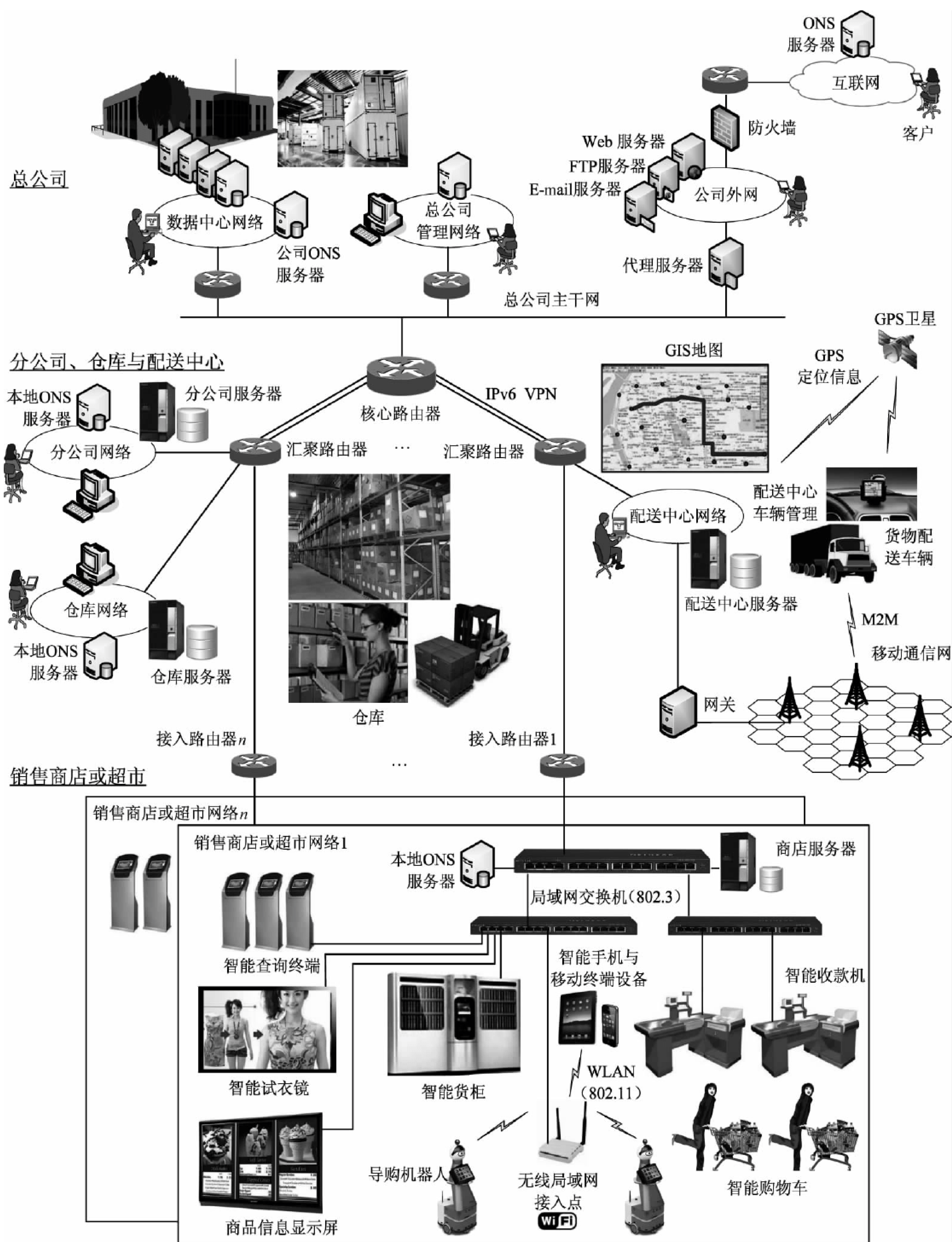


图 0-6　大型零售企业物联网网络系统结构示意图

路径与就近配货的通知。在制定不同地区的商品宣传、促销策略的过程中,需要使用海里数据存储与智能处理技术;在规划最佳配送方案,调度货物配送车辆的行车路径与就近配货的过程中,需要使用GPS定位、位置服务技术与智能控制技术。调度中心调度与货物配送车

辆直接需要通过移动通信实现信息的交互。支撑一个大型零售企业运行的计算与存储可能需要使用云计算平台,以及大数据分析模型与工具。

因此,设计、实现与运行一个大型零售企业物联网应用系统与一个基于互联网的应用系统,从应用的计算机网络技术与网络系统设计方法没有本质的区别。

0.2.4 从网络服务特点的角度去认识网络技术中的“变”与“不变”

如果我们从互联网与物联网所提供服务特点的角度作一个比较,我们就可以发现物联网与互联网技术的“变”与“不变”的关系。

(1) 互联网上传输的文本、视频、语音数据主要是通过人工方式产生的;物联网数据是通过RFID、传感器等感知手段,通过自动方式获取的。

(2) 互联网是虚拟的;物联网是虚拟与现实的结合。互联网上流传着一句话:“在互联网上没有人知道它是一条狗”。现在有人戏说:“在物联网上狗也是有‘身份’的网民”,这句话是正确的。它说明物联网是虚拟与现实的结合。

(3) 在互联网Web、FTP、E-mail,人们通过台式计算机、笔记本计算机、智能手机、PDA去发电子邮件、搜索信息、打电话、听音乐与看新闻,人与人之间的交互与信息共享都交给计算机来做;物联网的应用系统设计中将充分体现普适计算、CPS、可穿戴计算与环境智能化的设计思想,将计算机“装到”一切事物中,让人类享受更充分的“智能”服务。

(4) 互联网应用的设计者采取开放式的设计思想,试图建立面向全球客户的信息交互与共享网络信息系统;物联网设计思路是不同的,如智能工业、智能农业、智能电网、智能交通、智能物流等九大行业的应用,物联网提供的是行业性、专业性与区域性的服务。

(5) 互联网的电子邮件、文件传输、万维网、搜索引擎、即时通信、网络音乐、网络视频服务,以及到移动互联网应用为人类构建了一个人与人信息交互与共享的信息世界;物联网是通过“泛在感知、可靠传输、智慧处理”,最终将实现信息世界与物理世界的融合。

(6) 互联网应用为人类构建了一个人与人信息交互与共享的信息世界;物联网通过感知、传输与智能信息处理,生成智慧处理策略,再通过控制终端设备其实现对物理世界中对象进行控制。互联网与物联网最大的区别是:互联网是开环的信息服务系统,物联网系统是闭环控制系统。

互联网提供开环的信息服务系统与物联网提高闭环控制的不同,一定会对传输网的实时性、可靠性、安全性的要求上有一个大的变化。

0.2.5 从进程通信与TCP协议实现的角度去认识网络技术中的“变”与“不变”

在互联网应用层协议工作原理时都是做了一个假设:在一次进程通信过程中,源端与目的端之间一定要保证“持续”的TCP传输连接。如果不能保证TCP连接“持续”,那么分布式进程通信失败,网络服务不能实现。我们大量使用的互联网服务,如Web、E-mail、FTP都是建立在这个“假设”的基础之上。而目前存在着很多物联网应用,如水下无线传感器网络、地下无线传感器网络、军用无线传感器网络、GPS网络、无线车载网VANET、低地球轨道卫星通信网、星际网络等网络应用,实际上是运行在一个复杂的“受限网络(Challenged network)”之上,这个“假设”都是无法保证的。解决“受限”问题推动“容迟网(Delay

Tolerant Network,DTN)”技术的研究与发展。很多物联网应用系统都存在着“长延时、间歇性连接、低信噪比和高误码率、不对称数据速率与节点资源限制”等“受限”问题。

在一辆经常往返的公共汽车上安装射频通信系统就可以被用作信息存储和转发的工具。当这辆公共汽车从一个地方开到另一个地方时,它可以在附近的客户机和它将要去的地方的远程客户机之间提供信息交换服务。当地面移动节点与基站之间有建筑物阻挡,当地空卫星移动出卫星地面站接收范围,当无线车载网的节点之间收到其他车辆阻挡时,都会造成节点之间端-端连接的间歇性断开。这些端-端连接的中断可以有一定规律,也可以是随机的。但是,传统的TCP协议是不支持的。

无线、移动与长距离传输会导致接收信号的低信噪比与高误码率。在Internet中光纤传输的误码率可以达到10^{-12}～10^{-15},而太空通信中的误码率甚至可以达到10^{-1}。这种低信噪比、高误码率会极大地影响接收端对信号的解码和恢复,造成TCP连接的非正常中断,使得网络系统不能正常工作。

应用于太空、水下、战场、救灾现场与环境监测等环境中的无线传感器节点受体积和重量的限制,电源与计算、存储资源非常有限,它不可能像办公环境中有电源供应保障的PC一样,有足够的电源、计算与存储资源与网络带宽。无线传感器节点经常会因电池能量的耗尽而停止工作,其他节点要重新计算路由。无线传感器节点经常会因为节省电能而处于休眠状态,这是只有其他的节点唤醒它或休眠时间结束,它才能进入加入到无线自组网中的状态。在这种情况下,端-端连接也会经常中断。

在互联网中,传播时间一般以毫秒计算,而物联网中很多应用的数据传播延时会大大超出这个限度,属于长延时的应用,传统的TCP/IP协议是无法适应的。因此,研究物联网应用时必须讨论:在一次进程通信过程中,源端与目的端之间一定要保证“持续”的TCP传输连接。如何解决不能保证TCP“持续”连接下分布式进程通信的实现技术,是“容迟网(DTN)”技术研究的重点,也是物联网与互联网最重要的“变”数之一。

综上所述,在明确了互联网、移动互联网、物联网在技术与应用上的“变”与“不变”的问题基础上,我们需要对计算机网络课程中每个知识点“变”与“不变”的关系进行分析,研究教学内容中的“增”与“减”的关系。

我们采取的基本思路是:保留“不变”的核心技术,减少或删除已经“变”了的过渡性技术与淘汰的技术,适度增加新技术内容。

实事求是地说,我们对物联网技术与应用的研究不够深入,很多深层次的问题需要我们在今后的研究与工程实践继续研究和认识。

0.3 计算机网络技术发展对网络课程教学要求的变化

目前我国政府提出的战略性新兴产业、物联网、三网融合、大数据与云计算技术都是建立在计算机网络和互联网发展的基础之上。随着计算机网络技术的广泛应用,网络教育也开始由普及阶段向“扁平化”“多层次”与“专门化”方向发展。

0.3.1 扁平化

计算机网络本身是一门交叉学科,同时它也正在与其他的学科相结合,促进着新的

交叉学科的形成与发展。新一代信息技术、云计算、三网融合与物联网技术的发展,给计算机网络教育提出了重大的研究课题,网络课程的教学正在从计算机专业向相关专业发展,正在促进新的学科与课程的交叉。计算机网络已经成为软件编程的基本环境,所有学习计算机科学与技术、软件工程专业、网络工程、物联网专业的学生都需要学习和掌握网络环境中编程技术。计算机网络已经成为计算机及相关专业学生学习的一门基础性的课程。

0.3.2 多层次

现代信息服务业、软件产业,以及物联网、大数据与云计算技术与产业的发展,急需大量网络系统建设、应用软件研发、网络系统运维与管理的人才。无论是工科、理科,甚至是文科和艺术类,如网络工程、物联网、软件工程、电子商务、电子政务、物流、媒体传播、游戏制作、平面设计、广告等专业,很多课程的学习都是建立在学生掌握了网络知识和应用技术的基础上的。因此,不同学校、不同专业的教师都需要认真结合本专业的培养要求、办学特色,认真研究适应本专业教学要求的计算机网络课程的教学问题。

0.3.3 专门化

我国信息技术的发展和社会信息化程度的提升,使得社会对网络的依赖程度越来越高,这也导致了社会对网络人才需求的增加。目前各个单位急需大量网络系统组建、管理和维护人才。同时,Internet 应用技术、无线网络技术、物联网与网络安全技术领域也都离不开网络软件编程技术。即使是硬件系统的设计与实现,也会涉及大量的应用软件与嵌入式软件问题,这些都需要研发人员具备网络软件编程能力。社会急需大量网络应用软件与网络安全技术研发的专门人才。

图 0-7 给出了社会对计算机网络教学要求变化的示意图。



图 0-7　社会对网络教学要求的变化

0.4　从教学研究角度认识计算机网络课程改革的方向

0.4.1　以教学研究指导教学与教材体系建设

科学研究可以使我们从更高的角度、全局的视野去认识和把握技术的发展趋势，而教学研究对高质量的教学与教材体系建设能够起到重要的指导作用。

教学研究需要完成的工作任务是：与国内外同行专家进行学术交流，共同探讨提高网络课程教学质量的方法；浏览和跟踪各个学校的教学网站，研究和了解国内外知名大学网络课程教学内容、教材与主要参考书，以及作业与实验、教学方法与教学过程控制改革的动向；选择国际上最流行的教材，了解这些教材是如何处理新的技术发展与教学内容的关系。

有了国际知名大学的课程教学安排与教材作为参照物，借鉴美国 ACM 和 IEEE/CS 最新制定的计算机学科课程体系，才能够使得我们研究的网络教学与教材体系的内容、质量有可比性和评价依据。同时，通过对大型网络设备制造商与软件公司在员工培训、认证考试内容的变化情况的跟踪，了解产业界对人才需求的变化，结合作者在科研预研工作中对当前热点问题的研究、技术发展的了解，以及跟踪国内外知名大学每年教学内容与课程训练的更新情况，借鉴美国 ACM 和 IEEE/CS 最新制定的计算机学科课程体系，同时作者也认真分析了计算机专业硕士研究生入学统考大纲，不断调整、完善课程教学教材体系的总体设计思路与方案。这些工作成果为设计网络课程教学与教材体系提供了重要的决策依据。

0.4.2　研究计算机网络课程内容“变”与“不变”的关系

我们分析和比较的国际知名大学使用的教材主要有：Andrew S. Tanenbaum *Computer Networks*、Douglas E. Comer *Computer Networks and Internets with Internet applications*、James F. Kurose 与 Keith W. Ross *Computer Networking A Top-Down Approach Featuring the Internet* 等。

我们在细致地分析了 3 种教材的内容选取、章节组织、课程体系、作业与练习、参考文献等内容之后，对照图 0-8 所示的 Internet 发展的趋势图，其横坐标是时间，纵坐标是接入 Internet 的主机数量。*Computer Networks* 的 5 个版本分别出版于 1980 年、1988 年、1996 年、2003 年与 2011 年。在比对计算机网络与 Internet 技术发展过程的大背景之下，重点研究了 Andrew S. Tanenbaum 的 *Computer Networks* 第 1 版到第 5 版的内容更新过程。

Computer Networks 的第 1 版出版于 1980 年，当时网络技术研究处于开始阶段，因此教材的主要内容仅仅只是讨论了网络的一些基本概念。第 2 版出版于 1988 年，当时网络的使用还只限于大学与大公司，计算机网络在理论体系上处于形成阶段，OSI 参考模型与 TCP/IP 两大体系竞争激烈，因此教材的主要内容形成以介绍 OSI 参考模型为主的格局是很自然的事。第 3 版出版于 1996 年，当时正处于 Internet 开始广泛应用的阶段，因此教材以大量的篇幅讨论 TCP/IP 协议与 Internet 是符合当时技术发展水平的。第 4 版出版于 2003 年，当时的网络技术主要讨论的是 Internet 应用、高速局域网、无线局域网、网络安全，以及接入网技术，因此教材的第 4 版在概念与内容选取上一定会做比较大的调整。第 5 版出版于 2011 年，从第 4 版到第 5 版经历了 8 年。在这段时间内，Internet 在大规模接入的背

景下,促进了无线网络、移动IP、高速网络、三网融合技术的发展;在Internet应用广泛应用的背景下出现了基于C/S、B/S与P2P应用模式的演变,网络电话、网络视频、网络游戏、博客、即时通信、搜索引擎技术快速发展;在人与人交互技术日趋成熟的背景下,基于RFID、各种传感技术的人与物、物与物的物流网技术快速发展。这些应用背景带动了第5版内容的改变。为了解决Internet在IPTV等大型网络应用的QoS问题,出现了容迟网络(DTN)与内容分发网络(Content Delivery Networks,CDN)等新的研究方向。

图0-8给出了*Computer Networks*第3版至第5版出版的时间互联网发展状况对照图。认真研究*Computer Networks*第3版至第5版主要内容的演变,结合我们多年的教学经验,以及对计算机网络技术的理解,可以为我们把握计算机网络教学体系的设计、教材内容的选取具有重要的指导作用。

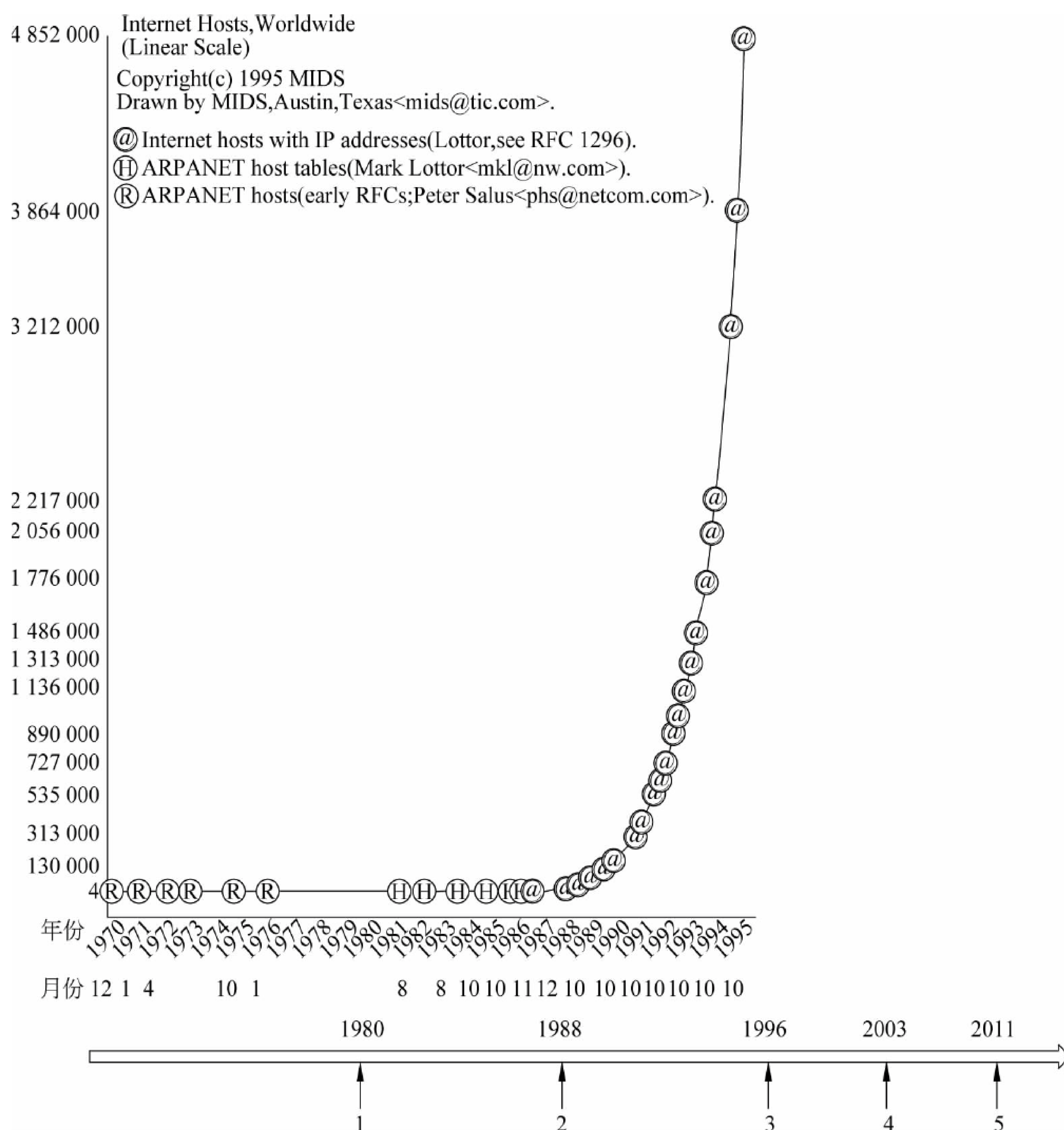


图0-8 *Computer Networks*第3版至第5版出版时间与互联网发展趋势对照

0.5 对 *Computer Networks* 第 3 版至第 5 版内容变化的分析

结合 *Computer Networks* 第 3 版至第 5 版的不同版本教材出版的时间，以及它们所处的互联网发展不同阶段的特点，我们就不难理解不同版本教材内容更新的背景。早期，我国计算机专业计算机网络课程教学的内容主要受 *Computer Networks* 第 1、第 2 版的影响，主要讲授 ISO/OSI 参考模型与七层协议的内容。*Computer Networks* 的第 3 版与前两版相比变化比较大，主要内容转移到讨论 TCP/IP 协议与互联网应用上。我们可以将对 *Computer Networks* 知识点变化的研究集中到第 3 版至第 5 版。需要注意的是，*Computer Networks* 第 3 版主要集中在 TCP/IP 协议与互联网应用上，第 4 版开始增加了无线网络的内容，第 5 版已经开始讨论物联网的相关技术。

0.5.1 第 1 章"计算机网络概论"内容的变化

"计算机网络概论"是对计算机网络技术的一个综述。研究"计算机网络概论"内容的变化可以帮助我们从宏观的角度，了解计算机网络技术总的发展趋势。表 0-1 给出了第 3 版至第 5 版"计算机网络概论"内容变化的比较。表中"＋"号表示增加的内容，"－"号表示删节的内容，黑体字表示比较重要的内容。

表 0-1 第 3 版至第 5 版"计算机网络概论"内容变化的比较

章节内容 \ 版次		第 3 版	第 4 版	第 5 版
第 1 章 计算机网络概论	1.1 计算机网络的应用	服务于企业的网络、服务于公众的网络、社会问题	商业应用、家庭应用、移动应用、社会问题	商业应用、家庭应用、移动应用、社会问题
	1.2 网络硬件	局域网、城域网、广域网、无线网络、互联网	局域网、城域网、广域网、无线网络、＋家庭网络、互联网	＋个人区域网局域网、城域网、广域网、无线网络、家庭网络
	1.3 网络软件	协议层次、各层的设计、接口与服务、面向连接的服务与无连接服务、服务原语、服务与协议的关系	协议层次、各层的设计、接口与服务、面向连接的服务与无连接服务、服务原语、服务与协议的关系	协议层次、各层的设计、接口与服务、面向连接的服务与无连接服务、服务原语、服务与协议的关系
	1.4 参考模型	OSI 参考模型、TCP/IP 参考模型、参考模型的比较、参考模型的特点	OSI 参考模型、TCP/IP 参考模型、参考模型的比较、参考模型的特点	OSI 参考模型、TCP/IP 参考模型、参考模型的比较、参考模型的特点
	1.5 网络实例	NetWare、ARPANET、NSFNET、Internet、Gbps 试验台	Internet、＋WLAN、＋Ethernet、－X.25、－FR、－ATM、－ISDN	Internet、WLAN、＋3G、＋RFID、＋传感网
	1.6 网络标准化	电信界组织、国际标准界组织、Internet 标准领域最有影响的组织	电信界组织、国际标准界组织、Internet 标准领域最有影响的组织	电信界组织、国际标准界组织、Internet 标准领域最有影响的组织

1. 第4版与第3版的比较

(1) 网络应用:第4版增加了商务应用、家庭应用、移动应用等内容。

(2) 网络硬件:第4版增加了无线网络与家庭网络等内容。

(3) 网络软件:基本相同,没有很大的变化。

(4) 参考模型:基本相同,没有很大的变化。

(5) 网络实例:第4版减少了SMDS、X.25、ISDN与ATM等内容,而增加了Internet与无线局域网WLAN的内容。

(6) 网络标准化:基本相同,没有很大的变化。

2. 第5版与第4版的比较

(1) 在计算机网络分类中,第5版在保持广域网、城域网与局域网不变的前提下,增加了个人区域网(PAN),这点正反映出近年来计算机网络技术的一个重要的变化。

(2) 在网络实例中,第5版在增加了3G、RFID与Sensor Networks的内容,这一点与物联网技术研究的发展是同步的。

0.5.2 第2章"物理层"内容的变化

表0-2给出了第3版至第5版"物理层"内容变化的比较。

表0-2 第3版至第5版"物理层"内容变化的比较

章节内容 \ 版次		第3版	第4版	第5版
第2章 物理层	2.1 数据通信的理论基础	傅里叶分析、有限带宽的信号、信道的最大传输速率	傅里叶分析、有限带宽的信号、信道的最大传输速率	傅里叶分析、有限带宽的信号、信道的最大传输速率
	2.2 传输介质	磁介质、双绞线、基带同轴电缆、宽带同轴电缆、光纤	磁介质、双绞线、同轴电缆、光纤	磁介质、双绞线、同轴电缆、光纤、+动力线
	2.3 无线传输	电磁波谱、无线传输、微波传输、红外与毫米波传输	电磁波谱、无线传输、微波传输、红外与毫米波传输	电磁波谱、无线传输、微波传输、红外与毫米波传输
	2.4 电话系统/通信卫星	电话系统的结构、本地回路、主干与多路复用、交换	通信卫星:地球同步卫星、中间过道卫星、低轨道卫星、卫星与光纤	通信卫星:地球同步卫星、中间过道卫星、地轨道卫星、卫星与光纤
	2.5 窄宽ISDN	ISDN服务、体系结构、接口、发展展望	公共数字电话网:结构、本地回路、+ADSL、干线与多路复用、交换	数字调制与多路复用:基带传输、通带传输、FDM、TDM、CDMA
	2.6 宽带ISDN与ATM	虚电路与电路交换、ATM、ATM交换机	移动电话系统:1G、2G、3G	公共交换电话网络:结构、本地回路、ADSL、光纤、多路复用、交换
	2.7 蜂窝无线通信	寻呼系统、无线电话、模拟移动电话、数字移动电话、个人通信服务	有线电视:基于有线电视的Internet、频谱分配、电缆调制解调器	移动电话系统:1G、2G、3G
	2.8 卫星通信	地球同步卫星、低轨道卫星、卫星与光纤		有线电视:卫星电视转播系统、基于有线电视的Internet、频谱分配、电缆调制解调器、ADSL与有线电视接入的比较

1. 第4版与第3版的比较

(1) 数据通信理论基础：基本相同，没有很大的变化。

(2) 传输介质：基本相同，没有很大的变化。

(3) 无线传输：基本相同，没有很大的变化。

(4) 通信卫星：基本相同，没有很大的变化。

(5) 公共交换电话网络：第4版增加了ADSL，减少了ISDN、ATM等内容。

(6) 移动电话系统：第4版增加了1G、2G、3G，减少了Cellular Radio等内容。

(7) 有线电视：第4版增加了Internet over Cable等内容。

2. 第5版与第4版的比较

(1) 第5版在传输介质中增加了动力线传输的内容。照明的220伏电线已经进入每一家、每一个办公室，借助电力线传输计算机网络数据信号的技术日趋成熟，并且在家庭网络中已经出现应用。增加这个内容能够适应家庭网络的应用需求。

(2) 第5版进一步增加了移动电话3G、有线电视等涉及三网融合的内容。

0.5.3 第3章"数据链路层"内容的变化

表0-3给出了第3版至第5版"数据链路层"内容变化的比较。

表0-3 第3版至第5版"数据链路层"内容变化的比较

版次 章节内容		第3版	第4版	第5版
第3章 数据链路层	3.1 数据链路层设计	数据链路层提供的服务、成帧、差错控制、流量控制	数据链路层提供的服务、成帧、差错控制、流量控制	数据链路层提供的服务、成帧、差错控制、流量控制
	3.2 差错控制	纠错码、检错码	纠错码、检错码	纠错码、检错码
	3.3 基本数据链路协议	单工协议、单工停止-等待协议、有噪声信道的单工协议	单工协议、单工停止-等待协议、有噪声信道的单工协议	单工协议、单工停止-等待协议、有噪声信道的单工协议
	3.4 滑动窗口协议	1位滑动窗口协议、GNB协议、SR协议	1位滑动窗口协议、GNB协议、SR协议	1位滑动窗口协议、GNB协议、SR协议
	3.5 协议描述与验证	有限状态机模型、Petri网模型	有限状态机模型、Petri网模型	有限状态机模型、Petri网模型
	3.6 数据链路层示例	HDLC协议、SLIP与PPP协议、ATM数据链路层	HDLC协议、SLIP与PPP协议	PPP over SONET、ADSL协议集

1. 第4版与第3版的比较

(1) 数据链路层设计要点：基本相同，没有很大的变化。

(2) 差错检测与纠正：基本相同，没有很大的变化。

(3) 基本的数据链路层协议：基本相同，没有很大的变化。

(4) 滑动窗口协议：基本相同，没有很大的变化。

(5) 协议验证：基本相同，没有很大的变化。

(6) 数据链路层协议示：第4版保留了HDLC、PPP，减少了ATM DLL等内容。

2. 第5版与第4版的比较

第5版与第4版比较大的变化是：第5版增加了PPP over SONET与ADSL协议集的内容。

0.5.4 第4章“介质访问控制子层”内容的变化

表0-4给出了第3版至第5版“介质访问控制子层”内容变化的比较。

表0-4 第3版至第5版“介质访问控制子层”内容变化的比较

章节内容 \ 版次		第3版	第4版	第5版
第4章介质访问控制子层	4.1 信道分配问题	LAN与MAN的静态分配、动态分配	LAN与MAN的静态分配、动态分配	LAN与MAN的静态分配、动态分配
	4.2 多路访问协议	ALOHA、CSMA、无冲突协议、有限竞争协议、FDM协议、WLAN协议、数字蜂窝无线电	ALOHA、CSMA、无冲突协议、有限竞争协议、WDM协议、WLAN协议	ALOHA、CSMA、无冲突协议、有限竞争协议、WLAN协议
	4.3 802协议/Ethernet	802.3、802.4、802.5、802.6、802.2	Ethernet：MAC、Ethernet性能、交换式Ethernet、Fast Ethernet、GE、LLC	Ethernet：物理层、MAC子层、Ethernet性能、交换式Ethernet、Fast Ethernet、GE、10GE
	4.4 网桥/WLAN	网桥：802.X-802.Y的网桥、透明网桥、源路由网桥、802网桥比较、远程网桥	WLAN：802.11协议集、物理层、MAC子层、帧结构	WLAN：802.11协议集、物理层、MAC子层、帧结构
	4.5 高速LAN	FDDI、Fast Ethernet、HIPPI、光纤通道	宽带WLAN：802.16协议集、物理层、MAC子层、帧结构	有限状态机模型、Petri网模型
	4.6 卫星网/蓝牙	轮询法、ALOHA、FDM、TDM、CDMA	蓝牙：体系结构、协议集、无线电层、基带层、L2CAP层、帧结构	蓝牙：体系结构、应用、协议集、无线电层、链路层、帧结构
	4.7 数据链路层交换/RFID		802.X-802.Y的网桥、本地网络互联、生成树协议、远程网桥、网络互联设备比较、VLAN	RFID：体系结构、物理层、标识码层
	4.8 数据链路层交换			网桥应用、桥的自学习、生成树协议、网络互联设备比较、VLAN

1. 第4版与第3版的比较

(1) 信道分配问题：基本相同，没有很大的变化。

(2) 多路访问协议：第4版减少了Cellular Radio等内容。

(3) Ethernet：第4版减少了802.4、802.5、802.6，增加了交换式Ethernet、FE与GE，没有10GE的内容。

(4) WLAN：第4版增加了802.11标准的基本工作原理与协议的内容。

（5）宽带无线网络：第4版增加了802.16标准基本工作原理与协议的内容。

（6）蓝牙技术：增加了蓝牙技术基本工作原理与协议的内容。

（7）数据链路层交换：第4版增加了网桥与VLAN的内容。

2. 第5版与第4版的比较

第5版与第4版最大的变化是增加了RFID应用系统体系结构、物理层与标识码层内容。这部分内容对于理解物联网应用系统工作原理、结构设计非常重要。

0.5.5 第5章"网络层"内容的变化

表0-5给出了第3版至第5版"网络层"内容变化的比较。

表0-5 第3版至第5版"网络层"内容变化的比较

章节内容 \ 版次		第3版	第4版	第5版
第5章 网络层	5.1 网络层设计要点	网络层服务功能、网络结构、虚电路与数据报	存储转发交换、服务功能、无连接与面向连接服务、虚电路与数据报	存储转发交换、服务功能、无连接与面向连接服务、虚电路与数据报
	5.2 路由选择算法	最优化原则、最短路由选择、扩散法、基于流量的路由、距离矢量路由、链路状态路由、分级路由、移动主机路由、广播路由、多播路由	最优化原则、最短路由选择、扩散法、基于流量的路由、距离矢量路由、链路状态路由、分级路由、广播路由、多播路由、移动主机路由、Ad Hod路由、P2P路由	最优化原则、最短路由选择、扩散法、基于流量的路由、距离矢量路由、链路状态路由、分级路由、广播路由、多播路由、任播路由、移动主机路由、Ad Hod路由
	5.3 拥塞控制算法	拥塞控制原理、预防策略、通信量整形、流说明、虚电路流控、抑制分组、延时抖动控制、多播拥塞控制	拥塞控制原理、预防策略、虚电路流控、数据报网络流控、负载丢弃、延时抖动控制	拥塞控制原理、基于流量的路由、进入控制、流量抑制、负载丢弃
	5.4 网络互联/QoS	无连接的网络互联、隧道、互联网路由选择、分段、防火墙	QoS：DiffServ、MPLS	QoS：DiffServ、MPLS
	5.5 IP协议/网络互联	IP协议、地址、子网、OSPF、BGP、多播、移动IP、CIDR、IPv6	网络互联：基本原理、隧道、网间路由、报分段	网络互联：基本原理、隧道、网间路由、报分段
	5.6 ATM网络层/IP协议	信元格式、连接建立、路由选择与交换、QoS、拥塞控制、ATM局域网	IP协议：IP协议内容、地址、子网、OSPF、BGP、多播、移动IP、CIDR、IPv6	IPv4协议、IP地址、IPv6、ICMP、MPLS、OSPF、BGP、IP多播、移动IP

1. 第4版与第3版的比较

（1）网络层设计要点：基本相同，没有很大的变化。

（2）路由算法：基本相同，没有很大的变化。

（3）拥塞控制算法：基本相同，没有很大的变化。

（4）服务质量：第4版增加了QoS、Diff-Service、MPLS等内容。

(5) 网络互联：第 4 版减少了防火墙的内容。

(6) Internet 网络层：第 4 版减少了 ATM NL 的内容，增加了 IP、ICMP、IGMP、OSPF、BGP、Mobil IP 与 IPv6 等内容。

2. 第 5 版与第 4 版的比较

第 5 版与第 4 版的比较变化不是很大，这说明网络层的内容相对比较稳定，第 4 版已经进行了详细地讨论。

0.5.6 第 6 章"传输层"内容的变化

表 0-6 给出了第 3 版至第 5 版"传输层"内容变化的比较。

表 0-6 第 3 版至第 5 版"传输层"内容变化的比较

章节内容 \ 版次		第 3 版	第 4 版	第 5 版
第 6 章 传输层	6.1 传输服务	传输层服务功能、QoS、传输服务原语	传输层服务功能、QoS、传输服务原语	传输层服务功能、QoS、传输服务原语
	6.2 传输协议要素	寻址、连接建立、连接释放、流量控制与缓冲策略、多路复用、崩溃恢复	寻址、连接建立、连接释放、流量控制与缓冲策略、多路复用、崩溃恢复	寻址、连接建立、连接释放、流量控制与缓冲策略、多路复用、崩溃恢复
	6.3 一个简单的传输协议	服务原语举例、传输实体举例、作为有限状态机的举例	服务原语举例、传输实体举例、作为有限状态机的举例	服务原语举例、传输实体举例、作为有限状态机的举例
	6.4 UDP/TCP 协议	TCP/UDP 协议：TCP 协议基本内容、TCP 连接控制、拥塞控制、传输策略、定时器管理、UDP 协议基本内容	UDP 协议基本内容	UDP 协议基本内容、远程调用、实时传输协议
	6.5 UDP/TCP 协议	ATM AAL 层协议	TCP/UDP 协议：TCP 协议基本内容、TCP 连接控制、拥塞控制、传输策略、定时器管理，UDP 协议基本内容、无线 TCP/UDP、事务型 TCP	TCP/UDP 协议：TCP 协议基本内容、TCP 连接控制、拥塞控制、传输策略、定时器管理、UDP 协议基本内容
	6.6 性能问题	网络性能概念、性能测试、优化性能的系统设计、适用于 Gbps 网络协议	网络性能概念、性能测试、优化性能的系统设计、适用于 Gbps 网络协议	网络性能概念、网络性能管理、高速长距离网络协议、优化性能的系统设计
	6.7 +DTN	网络性能概念、性能测试、优化性能的系统设计、适用于 Gbps 网络协议	网络性能概念、性能测试、优化性能的系统设计、适用于 Gbps 网络协议	+DTN 体系结构、DTN 协议

1. 第 4 版与第 3 版的比较

(1) 传输服务：基本相同，没有很大的变化。

(2) 传输协议要素：基本相同，没有很大的变化。

(3) 一个简单的传输协议：基本相同，没有很大的变化。

(4) UDP协议：基本相同，没有很大的变化。

(5) TCP协议：基本相同，没有很大的变化。

(6) 性能问题：第4版增加了GE性能的内容，减少了ATM AAL Layer protocol的内容。

2. 第5版与第4版的比较

第5版最大的变化是增加了6.7节，介绍了容迟网络(Delay-Tolerant Networking, DTN)的基本概念，讨论了DTN的体系结构与协议。容迟网络DTN技术的讨论标志着作者依据开始关注物联网的研究工作。

0.5.7 第7章"应用层"内容的变化

表0-7给出了第3版至第5版"应用层"内容变化的比较。

表0-7 第3版至第5版"应用层"内容变化的比较

章节内容 \ 版次		第3版	第4版	第5版
第7章 应用层	7.1 网络应用安全性/DNS	密码体系、对称加密、公钥加密、数字签名、社会问题	DNS：DNS名字空间、资源记录、名字服务器	DNS：DNS名字空间、资源记录、名字服务器
	7.2 DNS/SNMP	DNS：DNS名字空间、资源记录、名字服务器	E-mail：结构与服务、用户代理、消息格式、消息传输、最后的交付	E-mail：结构与服务、用户代理、消息格式、消息传输、最后的交付
	7.3 SNMP/Web	SNMP：模型、ASN.1、SMI、SNMP协议	Web：结构、静态Web页、动态Web页、HTTP协议、增强性能、无线Web	Web：结构、静态Web页、动态Web页、HTTP协议、十搜索引擎、移动Web
	7.4 SMTP/多媒体/音频流与视频流	E-mail：结构与服务、用户代理、消息格式、消息传输、电子邮件的隐私	多媒体：数字视频、视频压缩、音频流、Internet电台、IP语音、Mbone	数字音频、数字视频、流存储介质、实时交互
	7.5 USERNET/P2P	USERNET：概念、实现方法		P2P：容量与Internet流量、服务器集群与Web代理、CDN、P2P

1. 第4版与第3版的比较

(1) DNS：基本相同，没有很大的变化。

(2) E-mail：基本相同，没有很大的变化。

(3) Web：基本相同，没有很大的变化。

(4) IP Phone：基本相同，没有很大的变化。

(5) IP Broadcast：基本相同，没有很大的变化。

(6) Mbone：基本相同，没有很大的变化。

(7) 第4版将 Network Security、SNMP 调整到第8章。

(8) 第4版删去了 USERNET NEWS。

2. 第5版与第4版的比较

第5版最大的变化是增加了搜索引擎、移动 Web、P2P、服务器集群与 Web 代理,以及内容分发网络(CDN)的内容。

0.5.8 第8章"网络安全"内容的变化

表0-8给出了第3版至第5版"应用层"内容变化的比较。

表0-8 第3版至第5版"应用层"内容变化的比较

章节内容 \ 版次		第3版	第4版	第5版
第8章 网络安全	8.1 密码体系		密码体系、对称加密、公钥加密、一次一密、两条基本的密码学原则	密码体系、对称加密、公钥加密、一次一密、两条基本的密码学原则
	8.2 对称加密算法		DES、AES、其他的加密算法、密码分析	DES、AES、其他的加密算法、密码分析
	8.3 公钥加密算法		RSA、其他的加密算法	RSA、其他的加密算法
	8.4 数字签名		对称数字签名、公钥数字签名、消息摘要、生日攻击	对称数字签名、公钥数字签名、消息摘要、生日攻击
	8.5 公钥管理		证书、X.509、PKI	证书、X.509、PKI
	8.6 通信安全		IPSec、防火墙、VPN、无线网络安全	IPSec、防火墙、VPN、无线网络安全
	8.7 认证协议		基于共享密钥的认证、基于 PKI 的认证协议、使用公钥密码的认证协议	基于共享密钥的认证、基于 PKI 的认证协议、使用公钥密码的认证协议
	8.8 电子邮件安全		PGP、PEM、S/MIME	PGP、S/MIME
	8.9 Web 安全		安全命名机制、SSL、移动代码的安全	安全命名机制、SSL、移动代码的安全
	8.10 社会问题		隐私、言论自由、版权	隐私、言论自由、版权

1. 第4版与第3版的比较

第8章是在第4版中增加的。第8章的主要内容包括:

(1) 密码学基础:密码学的发展历史与基本概念。

(2) 对称密钥算法:技术特征与基本工作原理。

(3) 公开密钥算法:技术特征与基本工作原理。

(4) 数字签名:技术特征与基本工作原理。

(5) 通信安全:基本概念与方法。

(6) 认证协议:基本概念与方法。

（7）E-mail 安全：基本概念与方法。

（8）Web 安全：基本概念与方法。

（9）社会问题：基本概念。

2. 第 5 版与第 4 版的比较

第 5 版与第 4 版在第 8 章"网络安全"总体变化不大。

总结：

在对 *Computer Networks* 第 3 版至第 5 版内容变化分析的基础上，我们可以得出以下几点结论。

（1）第 3 版至第 5 版都有的内容，应该说是计算机网络技术最基本的概念和技术，由这些内容可以构成了网络课程主干的框架。

（2）第 4 版与第 5 版减少了 ATM、ISDN、X.25、帧中继等内容，符合当前技术发展的趋势。

（3）第 5 版增加的无线网络、移动 IP、P2P、搜索引擎、移动 Web、P2P、服务器集群与 Web 代理、PPP over SONET 与 ADSL 协议集等内容，符合当前技术发展的趋势。

（4）第 5 版增加的 RFID、Sensor Network、3G 等内容，这点与物联网研究与运用的发展是同步的。

（5）在计算机网络分类中，第 5 版在保持广域网、城域网与局域网不变的前提下，增加了个人区域网（PAN）的分类是恰当的。

（6）第 5 版增加的 DTN 体系结构与协议、CDN 等内容反映了近期计算机网络技术与物联网研究的热点问题。

需要注意的是，*Computer Networks* 第 5 版是 2011 年出版的，从 2011 年至 2016 年，发展最快的是无线网络技术，尤其是 Wi-Fi 技术的广泛应用。因此，在 2016 年修订的《计算机网络（第 4 版）》，需要重点增加无线网络 Wi-Fi 与 WSN 技术的相关内容。

0.6 本科网络课程教学定位与教材体系建设方案的设计

0.6.1 本科网络课程教学定位

科学研究为我们把握计算机网络技术的发展奠定了基础，教学研究为我们在网络课程改革指明了方向。针对目前计算机与相关本科专业的实际教学情况，将课程教学目标定位为：以网络技术基本理论与基本方法为主干，充分反映网络技术的最新发展，培养学生掌握网络的基本知识、基本工作能力与继续学习能力。

0.6.2 计算机网络课程教学与教材体系建设

总结作者与教学科研团队多年来本科与研究生网络课程教学的经验，以及指导研究生论文与科研工作的体会，吸取国内外知名大学成功的经验，考虑到计算机与相关专业本科与研究生阶段的计算机网络课程的基本要求，课题组经过近 20 年的能力，合作完成了本科到研究生的网络课程教学与教材体系（如图 0-9 所示）。

本科教材体系由主教材、教师用书、习题解析与同步练习、网络实验指导书、网络软件编

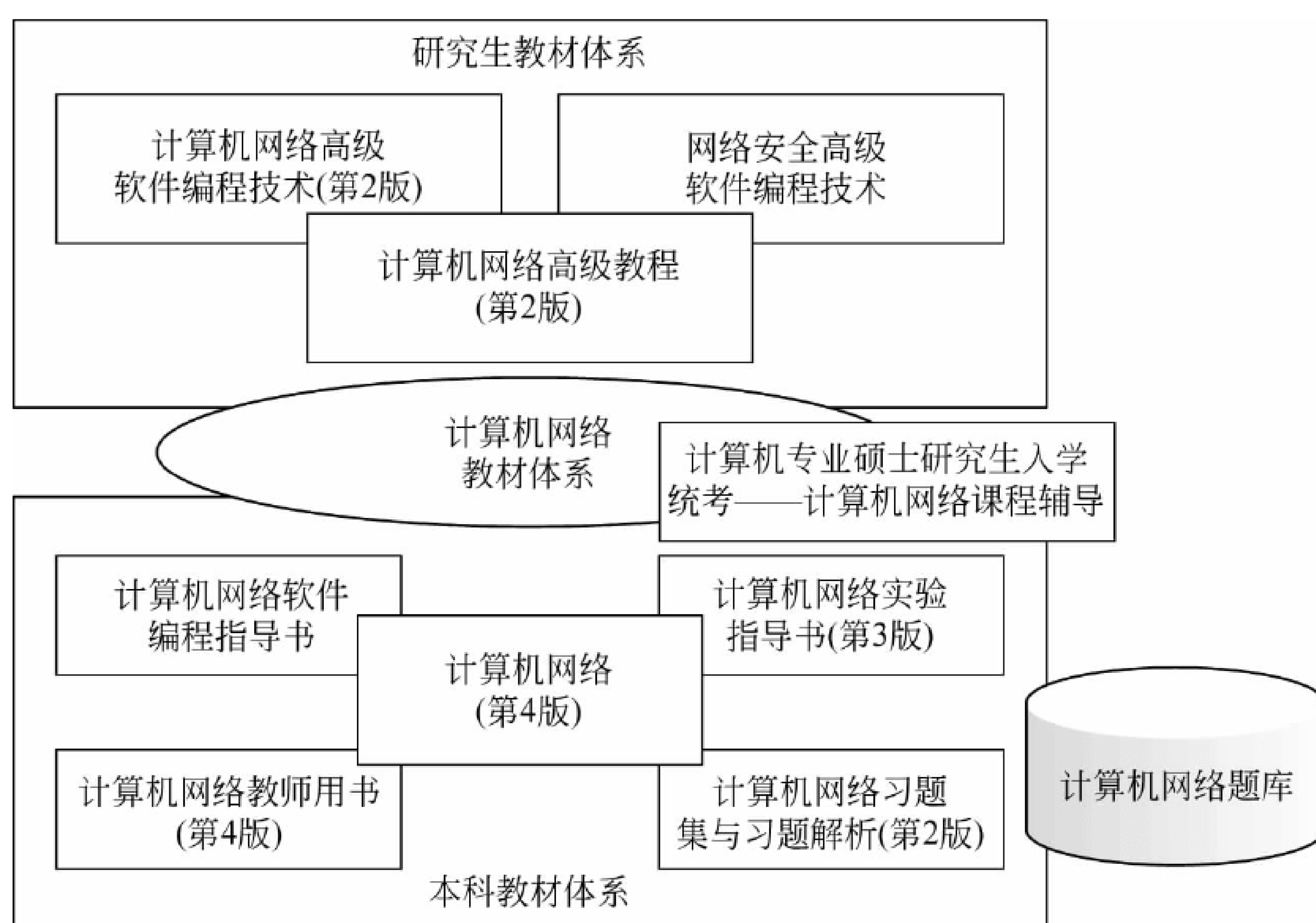


图 0-9　计算机网络课程教材体系

程指导书,以及与主教材配套的电子课件组成。其中,教师用书对教学方法与各章知识点的结构进行了系统的讨论,同时提供了部分关键技术发展的背景资料与协议标准查找方法。习题供学生自我检查知识掌握情况以及教师布置作业用。网络实验指导书、网络软件编程指导书分别可用于硬件实验与基本的网络软件编程训练。同时,由清华大学出版社组织,课题组完成了一个含有千道以上计算机网络习题题库的建设。

研究生计算机网络课程主教材是《计算机网络高级教程(第 2 版)》,配合理论课程的学习,提供了有数十个"近似实战"的网络软件编程与网络安全软件编程训练课题,分别出版了《计算机网络高级软件编程技术(第 2 版)》与《网络安全高级软件编程技术》。

由于我国设置有计算机及相关专业、学科点的学校很多,各个专业的师资、学生情况与培养目标等方面差异很大,不可能都要求统一配备有网络设备和硬件实验室。在这种情况下,由主教材来达到在宏观上把握教学基本要求,不同类型的学校可以根据各自的情况,自主地选择是通过硬件实验,还是通过软件编程的方法,或者两者相结合的方法,来达到提高学生实际工作能力培养的目的。这种思路应该是比较有效和符合国情的。

0.6.3　网络课程内容先进性与系统性的关系

多年的教学实践使我们认识到,在本科计算机网络课程教学与课程体系建设中需要解决好以下几个关系问题。第一个是计算机网络课程内容先进性与系统性的关系。

在网络课程的教学中,教师和同学都会遇到两个困难的问题。一是网络发展十分迅速,知识更新快,新的技术、新的术语不断出现。不要说是对于初次学习网络知识的同学,即使是多年从事网络技术研究与教学的专业人员也经常对技术的快速发展感到困惑。二是在网络课程教学过程中,讲授网络知识不讲网络体系结构与网络协议是不行的。但是如果完全按照传统网络的层次结构和协议去讲解,学生会感到枯燥和抽象,难以接受。如果不考虑网络体系结构与层次关系,教学中知识的组织会显得零散,读者很难掌握。要很好地解决这个

问题，教学体系的设计与教材的编写必须首先回答两个问题：一是教学内容的选择；二是教材体系的组织。

我国信息技术与产业的发展，需要大量从事计算机应用系统设计、网络系统集成工程师、软件工程师、电信技术、信息服务与各类信息系统管理的专业技术人员，以及网络与信息系统的使用和维护人员，他们都需要掌握网络知识与技术。计算机与通信是近年来两个发展最快、应用最广的学科，也是社会对人才需求最迫切的学科之一，而计算机网络正是这两个学科交叉发展的产物。因此，学习计算机网络必然会涉及计算机知识与通信知识两个方面的基础问题。从计算机科学与技术学科的角度，计算机网络是这个领域发展最为迅速的技术之一，也是计算机应用空前活跃的领域之一。如果从教学体系的角度，计算机网络的前期课程与基础应该包括计算机原理、操作系统、数据库技术等知识，以及基本的编程能力，同时深入研究计算机网络技术应该具备概率论、随机过程与排队论的基本知识。

计算机网络技术与通信技术相互渗透、密切结合而形成的一门交叉科学，它的内容必然要涉及通信技术的内容，在学习过程中会涉及一些通信技术方面的知识，计算机网络技术是建立在数据通信的基础之上。但是一定要注意，计算机网络课程不是数据通信课程。计算机网络课程是从计算机原理与应用的角度，去阐述计算机网络的基本工作原理与实现技术，要让学生学习和理解处理网络环境中软件、硬件与系统设计与实现方法，为深入掌握网络应用系统设计与编程知识打下基础。计算机专业的学生在计算机学科具有较为深入与广泛的知识基础，但是对通信技术知识基础不够，一般大学的计算机系也不会专门开设一门数据通信方面的课程，因此本书在编写中需要处理好：如何在知识点的组织中逐步加入有关通信技术的有关知识，而且学生只要具有物理学方面的基础，就能够接受这些知识，不需要学生专门去补这门功课。这是网络教材结构设计和教学过程组织中必须解决的一个重要问题。

0.6.4　网络理论教学与能力培养并重的关系

计算机网络是一门实践性很强的专业课程，如果教学值停留在理论探讨的阶段，学生考试之后可能只能够记住一些技术术语了，对于提高教学质量是远远不足的。计算机网络技术从 20 世纪 60 年代开始发展以来，已经形成了自身比较完善的体系。目前由于应用广泛，因此技术发展十分迅速，知识更新快，新的技术不断出现。对于这样一个发展迅速的领域来说，一本教材最重要的是让读者能够学会处理网络问题最基本的方法，掌握网络最基本的工作原理，使读者面对不断变化的技术，具有跟踪、学习的基础与能力。

目前本科毕业生在求职过程中反映出的实际动手能力弱的缺陷，与课程教学过程中的硬件实验与软件编程训练量的不足与要求不严有直接的关系。要提高教学质量，提高学生就业的竞争力，必须加强实践环节的训练。目前，网络课程教学急需解决理论与实际的结合，加强学生实际能力培养的问题。网络是一门应用性与实践性很强的课程。学生只有通过严格地实践训练，才有可能真正掌握和深入理解网络技术的基本理论、协议与算法。网络实际工作技能体现在网络系统规划、设计、组网与管理、维护，以及网络软件编程两大方面。针对这样的教学需求，教学团队编写了网络实验指导书、网络软件编程指导书等两本能力培养的辅助教材。

0.6.5 教材体系适用的范围

一套好的教材应该具有很好的适应性。目前我国设置有计算机专业的学校很多,师资与学生情况、培养目标等方面差异很大,不可能有统一的教学内容、学时的要求,也不可能都要求统一配备有网络设备或硬件实验室。在这种情况下,一种有效和符合国情的办法是:由主教材在宏观层面上把握教学基本要求,通过选择不同的辅助教材来适应不同类型、要求与条件的学校,达到提高学生实际工作能力培养的目的。针对本科学生中一部分将继续攻读研究生,而大部分学生将走向工作岗位的实际要求,在网络课程教学与教材体系的设计中必须针对不同需求的学生,提供相应的训练内容。

0.7 计算机网络课程教学内容

0.7.1 主教材《计算机网络(第4版)》知识点结构

1. 基本思路

通过近20年的教学实践,以及在与国内外同行的交流中,作者逐步地认识到:在本科教材的编写中,一定要以“不变”的网络基本概念、基本方法与核心技术为“主干”来组织章节,将“变化”的发展部分在适当的章节加入,替代已经过时的内容,以形成“主干”结构相对稳定,“变化”发展的新技术内容能够及时跟进的格局。

尽管网络课程所涉及的内容很多,但是和其他的学科一样,技术的发展必然有一个很自然的发展轨迹,这正体现出技术成熟的程度。面对快速发展的技术,我们只能从网络最基本的原理出发,总结提炼,让读者能够循序渐进地了解技术的发展过程,理解网络设计的基本思想。要做到这一点就需要处理好网络课程中“变”与“不变”的基本关系问题。解决这个问题需要有科研与教学研究经验与成果作为支撑。我们在分析 Andrew S. Tanenbaum *Computer Networks* 第4版与第3版中得出了这样几个结论:第4版保留下来的内容,应该说是网络技术在发展过程中一直“不变”的基本概念和核心技术;第4版变化的部分主要在 Internet 应用技术、无线网络技术与网络安全技术方面,这种修改正好顺应了网络技术的发展趋势。这些研究成果为我们在处理好课程内容的“变”与“不变”关系问题上起到了重要的指导作用。这也许是处理好网络课程教学中“面包”和“猎枪”的关系,以“不变”应“万变”的基本方法。

同时,我们需要注意的是:网络最重要的应用是 Internet。Internet 对科技进步和社会发展的影响是非常重大的,并且这种影响还在继续扩大,这也正体现出网络技术发展的生命力之所在。因此,学习网络知识重要的是通过学习网络基本知识,学会去认识 Internet 的核心技术,理解 Internet 的实现方法,希望通过本课程的学习,学生能够初步掌握与网络和 Internet 技术相关信息技术研究、应用与软件开发的知识,为今后的继续学习专业课程打下基础。同时,进入本课程学习的同学一般都具有使用 Internet 的经验,并且很多同学可能已经具有相当丰富的使用经验,因此本书以 Internet 技术为主线去组织知识点,有利于调动同学学习的积极性和主动性,引导学生从感性认识出发,结合基本理论的学习,逐步掌握网络技术中处理问题的基本方法,逐步培养将网络技术应用到各种行业的意识和能力。高速网

络技术最能够体现网络发展的方向与成果，以 Internet 技术与高速网络技术为主线组织本课程的知识体系，可以实现本课程希望达到的教学目标，也是国际知名大学的使用教材共同的特点。

本书知识点组织的基本思路是：遵循分层模型，但不拘泥于分层结构模型；以 Internet 技术与高速网络技术为主线，加入网络发展的最新成果。

2. 主教材结构

主教材的结构如图 0-10 所示。

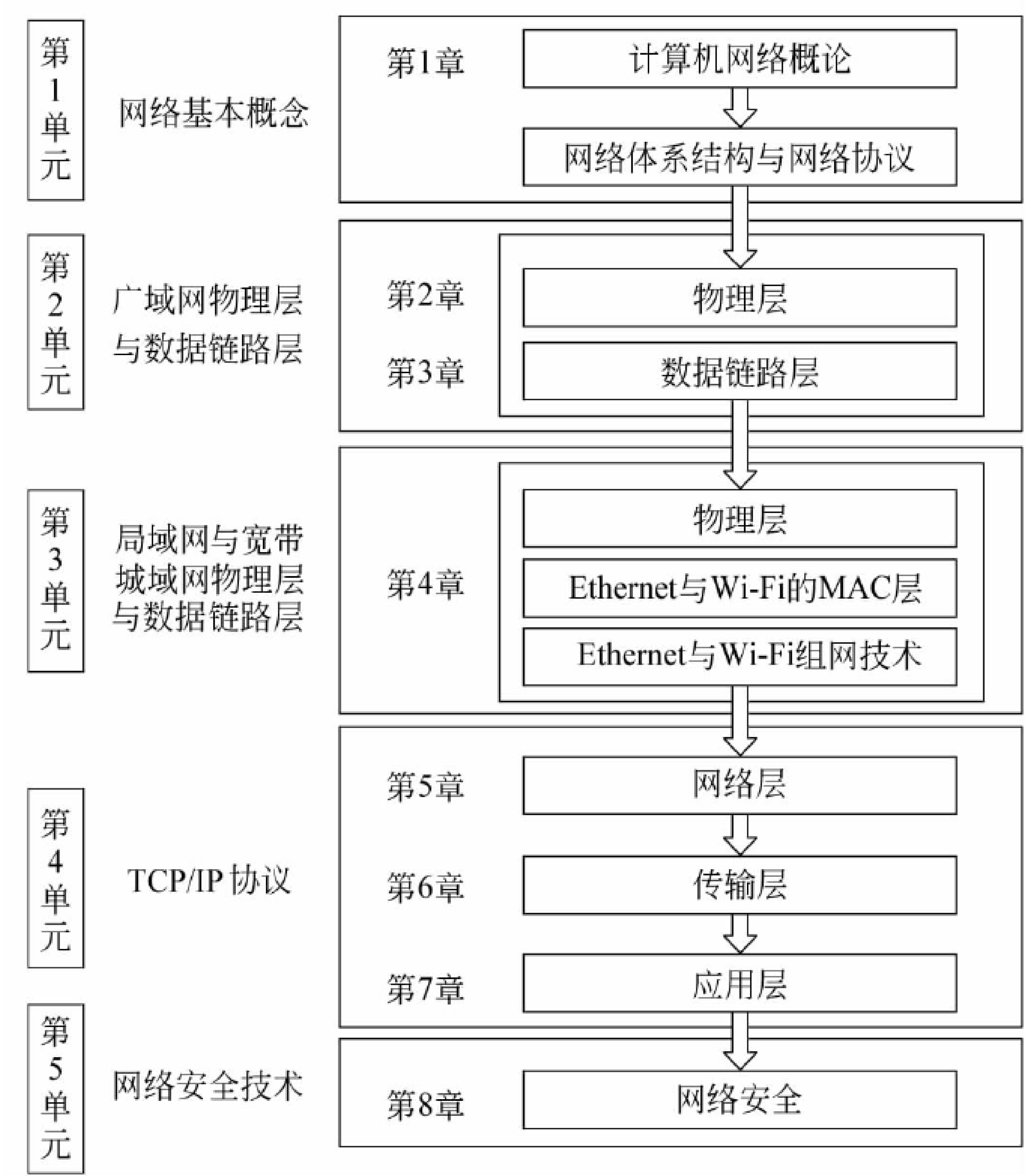


图 0-10 《计算机网络(第 4 版)》主教材结构

3. 关于广域网与局域网、城域网的物理层、数据链路层内容的安排

(1) 广域网与局域网、城域网技术之间的关系。

学习网络重要的问题之一是掌握广域网和局域网、城域网的基本工作原理。有些教材专门分出两到三章的内容，分别讨论了广域网和局域网、城域网技术。也有的教科书是按照物理层、数据链路层，将广域网和局域网、城域网放在一起，分在两章中讲。

本书采用了另外一种思路，那就是仍然保留分层的结构。由于广域网和局域网、城域网所采用的通信线路类型与数据交换机制的不同，它们之间有一些相同的地方，但是在技术特点上存在着较大的差异。因此，在网络的物理层和数据链路层协议上出现了两个分支，一类是基于点-点通信线路，另一类是基于广播信道。

基于点—点通信线路的广域网的物理层、数据链路层的技术与协议的研究开展得比较早，在局域网出现之前就已经初步形成了自己的体系、协议与标准。而基于广播信道的局域

网、城域网的物理层和数据链路层协议研究相对比较晚一些。人们研究了基于广播信道通信特点的体系、协议与标准。根据局域网的技术特点,对OSI参考模型中的数据链路层做了相应的修改,引入了介质访问控制子层与逻辑链路控制子层,在这些方面广域网和局域网存在着比较大的区别。但是在网络层以及高层,它们可以使用共同的协议。如果仅仅简单地根据物理层和数据链路层的划分角度,将两种体系的内容放在一起讨论,对于初学者来说掌握起来有一定的困难。这个问题可以用图0-11表示。

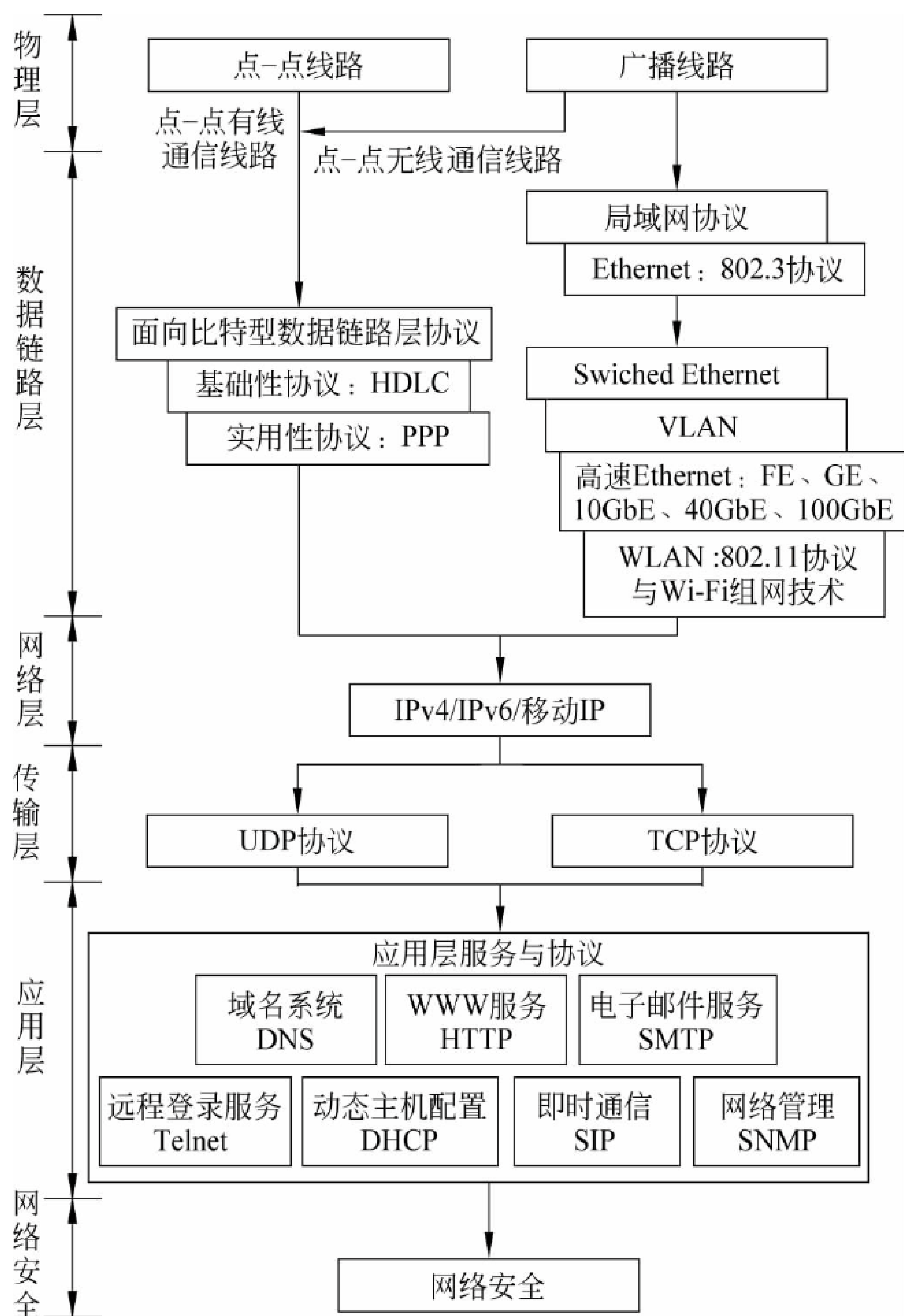


图0-11 网络协议演变的过程与相互关系示意图

从图0-11中可以看出,从物理层点-点通信线路与广播线路出发,在数据链路层形成了两个分支。

基于点-点通信线路的数据链路层协议的研究与应用经历了从面向字符型数据链路层协议向面向比特型协议的演变过程。所有数据链路层面向比特型协议的基础是HDLC,之后出现的很多实用的协议基本上都是在HDLC协议的基础上,取其子集而形成的。目前应用最为广泛的是PPP协议。这一条线所涉及的内容基本上已经成熟,因此协议没有太大的变化。

另一条线是基于广播线路的数据链路层协议,从网络分类的角度,它属于局域网一类。这一类是近年来发展最快的一支。局域网技术经过近20年的竞争,基本上形成了以

Ethernet 技术、IEEE 802.3 协议、MAC 层 CSMA/CD 控制技术为主体的局面。目前的发展主要体现在交互式 Ethernet、虚拟局域网、速率最高达到 100Gbps 的高速 Ethernet 等方面。同时,光以太网、城域以太网技术的发展使得 Ethernet 从局域网逐步扩展到宽带城域网与广域网。IEEE 802.11 协议与 Wi-Fi 无线局域网应用成为当前应用的热点。因此,这一部分内容属于数据链路层"变"的部分。

如何在网络教学中,以"不变"的技术为基础,形成一个相对稳定的教学框架,同时结合技术的发展,加入当前技术发展与应用热点的"变"的新技术,并能够使它们形成有机的整体,这是本教材重点要解决的问题,也是主教材的特色之一。

考虑到计算机专业学生前期学习的基础,以及技术发展的现状与趋势,为突出计算机专业培养的重点与教学特色,引导计算机专业学生用"系统观"去认识和理解网络原理与实现技术,主教材在内容选取上做了以下的调整:

- 适度地减少了物理层有线与无线通信技术细节的内容。
- 将计算机网络原理与实现技术的讨论,从互联网扩大到移动互联网与物联网。
- 增加了从计算机系统的角度认识、掌握和应用网络技术的内容。
- 提高了计算机网络协议软件编程能力训练的要求。

(2) 本书的处理方法。

在本书的结构设计中,第 2 章讨论基于点-点通信线路的物理层协议与标准,以及相关的数据通信技术;第 3 章讨论基于点-点通信线路的数据链路层协议与标准。第 3 章、第 4 章主要以广域网的物理层、数据链路层技术为背景。第 4 章以局域网、城域网与无线网络技术为背景,讨论基于广播信道的物理层和数据链路层协议与标准。在第 2 章、第 3 章与第 4 章的基础上,第 5 章、第 6 章讨论它们可以共同使用的网络层 IP 协议与 TCP 协议。这样的知识结构组织既符合技术发展的规律,也比较能够适应读者循序渐进学习的需要。这种知识结构的组织方法是根据我们多年的教学经验,以及根据我们在科学研究中对技术发展走向的理解基础上形成的。从本身实际教学结果以及使用本教材老师反馈回来的意见,这种知识结构的组织方法是比较合理的。

(3) 主教材知识点安排的内在关系。

从教材章节结构中可以看出,全书构成了一个清晰的结构。教材中的每一章内容呈逐步递进的关系。每一章内容回答了网络课程中一个基本的问题。这些问题是:

第 1 章　基本概念:什么是计算机网络?

第 2 章　物理层:网络中比特流的传输是如何实现的?

第 3 章　数据链路层:网络中数据传输的正确性是如何保证的?

第 4 章　介质访问子层:最常用的 Ethernet 与 Wi-Fi 的网络功能是如何实现的?

第 5 章　网络层:网络互联是如何实现的?

第 6 章　传输层:网络环境中分布式进程通信是如何实现的?

第 7 章　应用层:网络应用系统是如何设计和实现的?

第 8 章　网络安全:如何保证网络安全?

主教材知识点的内在关系如图 0-12 所示。

当前网络教材在写作语言上有一个突出的问题是翻译的痕迹重、概念表述含混和不严格。要提高学生学习的积极性与自主性,除了在体系、内容的选取上下功夫之外,教材写作

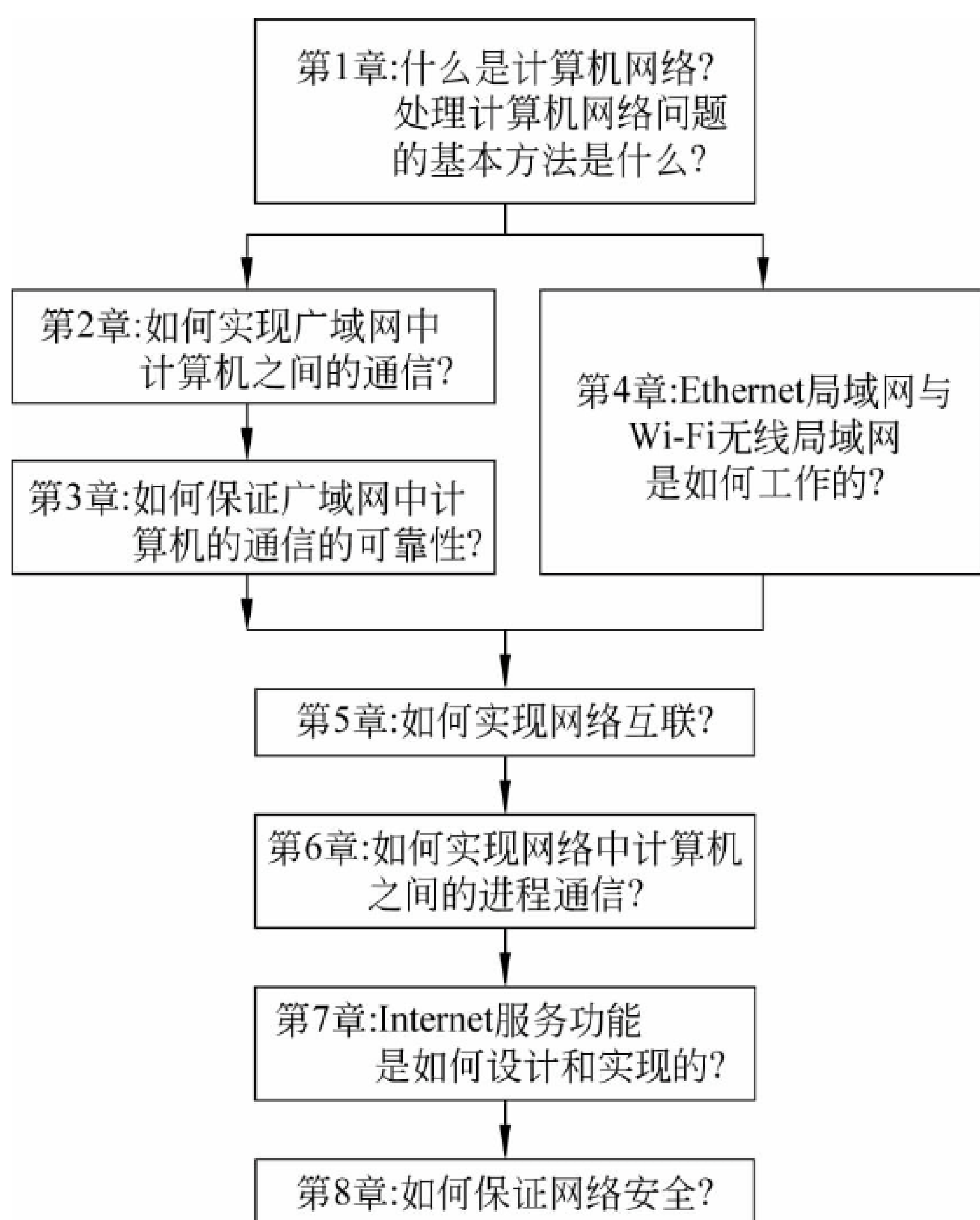


图 0-12　主教材知识点的内在关系

语言上要符合中文的语言习惯,由工作在教学、科研第一线的老师编写的好处是:贴近技术发展的前沿,可以用研究者自己的体会去解读技术的内涵。主教材的编写力求做到:

- 主干结构:清晰,重点突出。
- 前后内容:衔接,循序渐进。
- 语言表达:准确,易于自学。

0.7.2 《计算机网络教师用书(第4版)》的编写

教师用书希望在这一方面给初次教授网络课程的老师有所帮助。作者在编写这本教材时注意考虑了以下几个方面的问题:

(1) 教材的知识点是如何组织。

(2) 讲授这些内容时需要了解的一些背景知识。

(3) 希望学生掌握哪些知识。

(4) 哪些概念不容易理解,哪些问题容易混淆或容易出现错误。

(5) 主要的技术标准、文档资料来源与重要网站。

作者在《计算机网络教师用书》中,作者根据多年教学科研工作的经验与体会,总结多年来教师与教师、教师与学生之间讨论的问题,结合作者自己在学习过程中不断发现的疑点、难点,按照主教材的章节结构顺序,以节为单元,通过提出了300个问题,以提问与回答的方式,解析技术内涵与背景知识,剖析容易产生歧义问题的原因,帮助初次承担网络课程教学的老师,熟悉教学内容、细节与难点。

0.7.3 《计算机网络实验指导书》的编写

高质量的网络实验与实验教材是目前亟待解决的问题。团队成员总结多年从事网络系统规划与设计、组建与管理的实践经验与科研工作的体会，以及多年网络课程与实验教学经验，参考了国内外知名大学的实验和 Cisco 认证所涉及的内容，构思了 13 个网络实验题目。设计的实验内容包括了从物理层的数据传输到应用层的网络应用，同时还涵盖了网络安全与网络仿真程序应用等主要的内容。实验指导书在每个实验内容之后，给出了进一步掌握该实验内容的练习与思考题。网络实验的题目为：

实验 1：简单的异步串行通信编程实验

实验 2：利用停止等待协议传输数据文件

实验 3：Ethernet 网组网实验

实验 4：虚拟局域网 VLAN 的配置

实验 5：网络数据包的监听与分析

实验 6：获取以太网中 IP 地址与 MAC 地址的对应关系

实验 7：路由配置和简单的路由程序

实验 8：编写简单的客户/服务器程序

实验 9：域名服务器的配置

实验 10：编写简化的 SMTP 邮件服务器并观察其通信过程

实验 11：Web 服务器的配置和管理

实验 12：利用 SSL 实现安全数据传输

实验 13：网络仿真

实验所要求的设备相对比较基本和简单，目前很多学校的网络实验室基本上都具有这些条件。

0.7.4 《计算机网络软件编程指导书》的编写

设计《计算机网络软件编程指导书》的目的有两个。如果有的学校缺乏基本的网络实验教学条件，但是不希望降低对学生实际能力培养的要求，可以安排学生在普通的计算机机房的教学环境中，不需要添加特殊的网络设备，只要学过 C 或 C++ 语言的学生就可以完成的网络编程训练，以弥补网络实验硬件条件的不足。同时，教学科研团队成员在带本科毕业论文的过程中，发现很多计算机专业的本科学生编程能力不强，对网络软件编程更不入门。现代的软件都是运行在网络环境中，如果能将两者紧密地结合起来，让学生通过网络软件编程的训练过程去加深对网络理论的理解，同时也能够提高学生基本的软件编程能力。

基于这样的认识，本书作者参考国内外知名大学网络课程软件编程训练与著名信息技术企业员工网络软件编程培训中的相关资料与文献，构思了 13 个网络软件编程题目。网络软件编程题目的选择考虑到网络不同的协议层次，同时将编程题目分为 3 个难度级，读者可以参考选题指导，根据不同的要求和不同的基础，有选择地、循序渐进地完成网络软件编程训练，达到“通过实际编程问题的训练，达到加深理解网络基本工作原理，掌握网络环境中软件编程方法，提高网络软件编程能力”的目的。软件编程课题的练习目的、要求与难度级如表 0-9 所示。

表0-9 软件编程课题的练习目的、要求与难度级

序号	课 题	层次	练习目的	难度级别
1	Ethernet帧的封装与解析	数据链路层	① 掌握Ethernet帧结构中各字段的含义与用途。 ② 掌握Ethernet帧结构解析软件设计与编程方法。	★
2	Ethernet 帧的CRC校验		① 掌握CRC校验的原理与计算方法。 ② 理解数据链路层协议的设计思想、原理与编程方法。	★★
3	IP地址的合法性判断	网络层	① 掌握IPv4地址基本结构与分类方法。 ② 理解网络层协议的设计思想与工作原理。	★
4	IP分组的捕获与解析		① 掌握IP分组头各个字段的含义与用途。 ② 掌握通过网卡截获经过的IP包的基本方法。	★★
5	IP分组的分片与重组		① 掌握IP分组分片与重组的原理与方法。 ② 理解网络层与数据链路层、物理层之间的关系。	★
6	IPv6分组封装与解析		① 掌握IPv6分组头各个字段的含义与用途。 ② 理解IPv6协议的设计思想与工作原理。	★★
7	发现网络中活动主机		① 掌握ICMP分组各个字段的含义与用途。 ② 理解ICMP协议设计思想、原理与编程方法。	★★★
8	发现服务器开启的TCP端口	传输层	① 理解传输层分布式进程通信与端口的概念。 ② 掌握端口扫描的工作原理与编程方法。	★
9	TCP报文的封装与发送		① 掌握TCP报文头各字段的含义与用途。 ② 理解TCP协议的设计思想与工作原理。	★
10	基于TCP的客户/服务器程序		① 理解TCP服务的特点与主要功能。 ② 掌握基于TCP的客户/服务器程序设计方法。	★★
11	基于UDP的客户机/服务器程序		① 理解UDP服务的特点与主要功能。 ② 掌握基于UDP的客户机/服务器程序设计方法。	★★
12	FTP客户机程序设计	应用层	① 掌握FTP服务的基本工作原理。 ② 掌握应用层协议的基本设计思路与编程方法。	★★★
13	包过滤防火墙程序设计		① 理解防火墙的基本概念与主要功能。 ② 掌握包过滤技术的设计思路与编程方法。	★★

0.7.5 《计算机网络习题解析与同步练习(第2版)》的编写

1. 习题解析与同步练习的目的

为了帮助学生在学习过程中对重要概念、知识理解与掌握情况进行的自我检查,也便于教师安排课后作业,作者参考了国外大学的作业与教材的习题,参考了Cisco与微软认证以及网络工程师、网络管理员、应聘等考试的内容,精选了一部分习题进行了解析,帮助同学循序渐进地理解知识。希望通过完成同步练习,帮助学生检查对网络基本概念理解的情况,为学生的自主学习创造条件。

2. 日常训练与硕士研究生入学统考的关系

计算机专业硕士研究生入学统考是很多本科学生毕业时必须面对的问题。计算机网络课程已经列入计算机专业硕士研究生统考的科目。但是从实施的效果来看,学生在比较短的时间内要面对数据结构、计算机组成原理、操作系统与计算机网络等4门课程的复习是很困难的。尤其是计算机网络课程,内容庞杂,而占的分数不多,因此很多考生实际上是放弃

了这门课程，疲于奔命地应试其他的三门功课，这对于日渐重要的计算机网络课程的日常教学是很不利的。

从考查目标可以看出，大纲要求考生通过网络课程的学习，掌握网络的基本工作原理，具有网络系统分析、设计与应用的基本能力。这是一个基本的考核标准，也是命题的指导思想与原则。这与学生在校学习计算机网络课程的要求应该是一致的。

习题解析与同步训练按照主教材的章节顺序安排。在这本习题解析与同步练习辅助教材没有采取传统的写作习惯，先在每一章的开始将整个一章的内容总结一遍，然后再进入习题解析部分。这样的辅助教材的内容重叠和庞大，复习起来困难。《计算机网络习题解析与同步练习》采取提供习题解析的方法，将每一章节重点的内容串起来，用尽可能小的习题量来覆盖需要复习和掌握的基本概念、知识和技能。大纲公布的试题示例中表示试题类型只有两种：单选题与综合应用题。例题与练习题的命题参考了大纲公布的命题格式，这样在完成例题与练习的过程也可以进一步地适应考试。

对于很多在网络知识上掌握得比较好的同学，他们可以通过快速地浏览例题和完成练习题，复习知识，找出不足，有重点地提高。对于网络知识上掌握得一般的同学，他们可能心里没底，不知道如何复习。建议这一类同学跟着辅导教材，不要先看例题解析，而是自己先做，然后与例题解析的过程去比较，在比较过程中发现自己知识点的缺陷，即时补充。然后通过独立完成练习题去检查自己的掌握情况，进一步找出不足，第二遍有重点的复习，争取能够在不太长的时间取得比较好的进步。因此希望读者能够在对于例题和做习题一样，先自己完成，做不出来再看解析，找出自己在哪些知识上需要加强，哪些错误的概念需要纠正；做出来的可以与解析比较，看看哪种方法更好。在完成例题与习题的过程中，再结合课本进行复习，这样会起到事半功倍的效果。

0.7.6 网络课程教材的使用与教学方法的讨论

由“一本主教材、四本辅助教材、一个题库和一个网络课件”构成的计算机网络课程立体教学体系为网络课程教学方法的改革提供了一个平台和基础。不同的学校可以根据课程教学定位、学时与设备条件，有选择地利用以上教学资源，开展自身的课程建设与教学方法的改革。

1. 教材应用模式

不同的学校应该有不同的重点和教学模式。以下是几种教材应用模式。

(1) 模式1。

《计算机网络》+《计算机网络教师用书》+《计算机网络电子教案》

(2) 模式2。

《计算机网络》+《计算机网络教师用书》+《计算机网络电子教案》+《计算机网络习题解析与同步练习》

(3) 模式3。

《计算机网络》+《计算机网络教师用书》+《计算机网络电子教案》+《计算机网络习题解析与同步练习》+《计算机网络实验指导书》

(4) 模式4。

《计算机网络》+《计算机网络教师用书》+《计算机网络电子教案》+《计算机网络习题

解析与同步练习》+《计算机网络软件编程指导书》

(5) 模式5。

《计算机网络》+《计算机网络教师用书》+《计算机网络电子教案》+《计算机网络习题解析与同步练习》+《计算机网络实验指导书》+《计算机网络软件编程指导书》

2. 讨论

(1) 模式1是一个基本结构,教学以《计算机网络》为主,使用主教材中每一章之后的习题来检查学生的学习情况;《计算机网络电子教案》《计算机网络教师用书》作为教师教学的参考。该模式适合学时安排比较少的学校教学需要。

(2) 模式2是一个基本结构,教学以《计算机网络》为主,在主教材中每一章之后的习题不够的情况下,可以使用《计算机网络习题解析与同步练习》来检查学生的理论知识学习情况;《计算机网络电子教案》《计算机网络教师用书》作为教师教学的参考。该模式适合学时安排比较少,但是理论学习要求相当高一些的学科教学的需要。

(3) 模式3是一个增强实验能力要求的结构,教学以《计算机网络》为主,使用《计算机网络习题解析与同步练习》来检查学生的理论知识学习情况;《计算机网络电子教案》《计算机网络教师用书》作为教师教学的参考。《计算机网络实验指导书》作为实验教学用书使用。该模式适合具有基本的网络实验条件,对于网络课程学习要求相当较高的学科教学需要。

(4) 模式4是一个增强软件编程能力训练要求的结构,教学以《计算机网络》为主,使用《计算机网络习题解析与同步练习》来检查学生的理论知识学习情况;《计算机网络电子教案》《计算机网络教师用书》作为教师教学的参考。在适当的章节之后,根据《计算机网络软件编程指导书》来安排学生在网络环境的编程训练。通过网络环境的编程训练,加强学生对网络知识的理解,提高学生在网络环境中的编程能力。该模式适合对于网络知识的理论学习和软件编程要求相对较高的学科教学需要。

(5) 模式5是一个全面训练的结构,教学以《计算机网络》为主,使用《计算机网络习题解析与同步练习》来检查学生的理论知识学习情况;《计算机网络电子教案》《计算机网络教师用书》作为教师教学的参考。《计算机网络实验指导书》作为实验教学用书使用。在适当的章节之后,根据《计算机网络软件编程指导书》来安排学生在网络环境的编程训练。通过网络环境的软件编程训练,加强学生对网络知识的理解,提高学生在网络软件编程能力。该模式适合对于网络知识的理论学习和软件编程要求相对较高的学科教学需要。

这样多种模式的设计是为了适应不同学校、不同专业对计算机网络课程学习的需要,便于教师根据教学计划的要求、学时的多少、学生的接受情况、教学条件等因素,以提高教学质量为目的,灵活地掌握和选择。

0.8 教材内容与研究生入学统考(网络技术)大纲内容要求的关系

0.8.1 对研究生入学统考(网络技术)大纲内容的分析

全国计算机专业硕士研究生入学统考大纲中规定的课程有计算机网络技术。很多本科毕业生需要参加研究生入学统考。无论教师对有些问题存在不同的看法,但这都是我们无

法回避的现实。因此，任课教师只能从积极的方面来对待这个问题，在日常教学中注意处理好教学内容与国家规定的统一考试的关系。我们只能通过高水平的网络课程理论教学与严格的能力训练，让学生比较牢固地掌握计算机网络的基本理论与动手能力，以提高学生继续升学与就业的适应能力。

计算机专业全国硕士研究生入学统考大纲对计算机网络课程的考查目标有以下三点：

（1）掌握计算机网络的基本概念、基本原理和基本方法。

（2）掌握计算机网络的体系结构和典型网络协议，了解典型网络设备的组成和特点，理解典型网络设备的工作原理。

（3）能够运用计算机网络的基本概念、基本原理和基本方法进行网络系统的分析、设计和应用。

分析大纲列出考查范围可以看出，被列入考查的知识点能够反映当前技术发展与网络课程教学的基本要求，与目前国内外流行网络教材的内容与结构保持了很好地一致性。

我们重点将大纲所要求考查的知识点与《计算机网络（第2版）》以及相关的教材做一个比对，就可以清楚地看出这一点。这对学生了解考试要求与我们日常教学要求之间的联系，增强备考的信心是有益的。当然，考纲对知识点的组织与某一本教材的结构不可能都是完全一样的。但是需要注意教材的内容、深度与考纲的要求的联系与区别。这里有几个问题需要说明：

第一，有些知识点对于理解网络工作原理十分重要，在网络课程的教学中必须要求学生掌握，同时入学考试也要求考查该项内容。例如说分组交换的概念，大纲放在第二部分的物理层，而很多教材是放在第1章计算机网络概论中。这种情况可能多处出现，同学们需要根据自己在学习这门功课所使用教材的具体情况来安排复习。

第二，即使是同一个问题，不同的教材或者是不同的教师，在讲授的深度和要求上都会有区别的。作者在与不同的学校毕业学生的接触中很明显地能够感觉到，有些学校在网络理论教学上要求比较高，有些学校对网络规划、设计、组网技术方面训练有素，这种区别是非常正常的，也体现出不同学校的教学特色。但是从全国性的入学统考角度，它只能够从一个统一的要求出发去命题，这样就存在着一个命题教师与不同学校考生的相互适应问题。当然从备考的角度，学生更需要去经过复习的过程，来不断完善对大纲要求知识点内容的掌握，通过考试争取最大限度地反映出自己对知识与技能掌握的程度。因此，每一位准备参加考试的同学需要有一个良好的“调整、充实、提高”的心态。

第三，大纲表述越简单，命题的灵活性就越大。大部分学校的网络教学更多的精力是集中在基本理论与基本技能的训练上，对于综合训练普遍感到不足。但是，大纲要求要考查学生的综合应用能力。

例如IPv4地址问题，不同的教材所涉及的深度差异很大。从大纲角度，它只能够用“IPv4地址与NAT”去表述。但是，从2009年入学统考命题的网络部分的综合题看，实际上就是一个涉及对一个标准的C类IP地址做子网划分，同时考核了路由表生成与地址汇聚问题。这样的综合问题有利于对IP地址以及子网划分、路由器工作原理理解比较深入的同学，以及在本科教学中做过这一类练习或实验的同学。对于从来没有做过这类问题的同学，解决起来是有一些困难的。同样一个问题，最简单的办法是变化一些求解的形式，或者是改变一个IP地址或改变一个子网的地址数，就必须重新算一遍，并且会出现完全不同的

结果。也就是说,同一个类型的问题可以在多次考试中出现也不存在重复命题的问题。这对于从教多年的教师来说,是一件很平常的事。当然这样做也说明命题教师已经形成了一个确定的认识,那就是学习过网络课程的同学必须掌握这些基本和重要的知识与技能。现在 IPv4 地址已经耗尽,但是我们目前仍然要在很长的一段时间仍然要面对使用 IPv4 的问题,逐步转向 IPv4 向 IPv6 过渡,下一阶段的重点将转移到 IPv6,这是很自然的事。

如果有的同学发现这个问题之后,就花了很大精力去补这一块知识,从准备报考研究生同学备考的实际情况来说是不允许的。作者在认真研究了考题之后,思考了帮助同学复习的办法。在复习的例题与习题设计中考虑到基础知识的训练,同时也适当地设计了一些综合性的例题和习题,通过有选择的训练,提高同学综合解决问题的能力,希望能够在帮助学生适应大纲要求方面做一些努力。

0.8.2 计算机网络体系结构

大纲将考查的知识点划分为 6 个部分。大纲的第一部分是“计算机网络体系结构”。它又是由 2 个知识点组成。

1. 第一个知识点:计算机网络概述

计算机网络概述包括以下 3 个方面的内容:

(1) 计算机网络的概念、组成与功能。

(2) 计算机网络的分类。

(3) 计算机网络与 Internet 的发展历史。

(4) 计算机网络的标准化工作及相关组织。

《计算机网络》第 1 章的内容与大纲“计算机网络体系结构”的知识点要求与教材第 1、2 章之间的对应关系如图 0-13 所示。教材中第 1 章讨论了计算机网络的定义、分类、功能、组成与发展历史,以及分组交换的基本概念。只是将计算机网络的标准化工作与相关组织放在第 2 章中结合网络体系结构与协议中讨论了。因此复习第 1 章能够覆盖大纲计算机网络

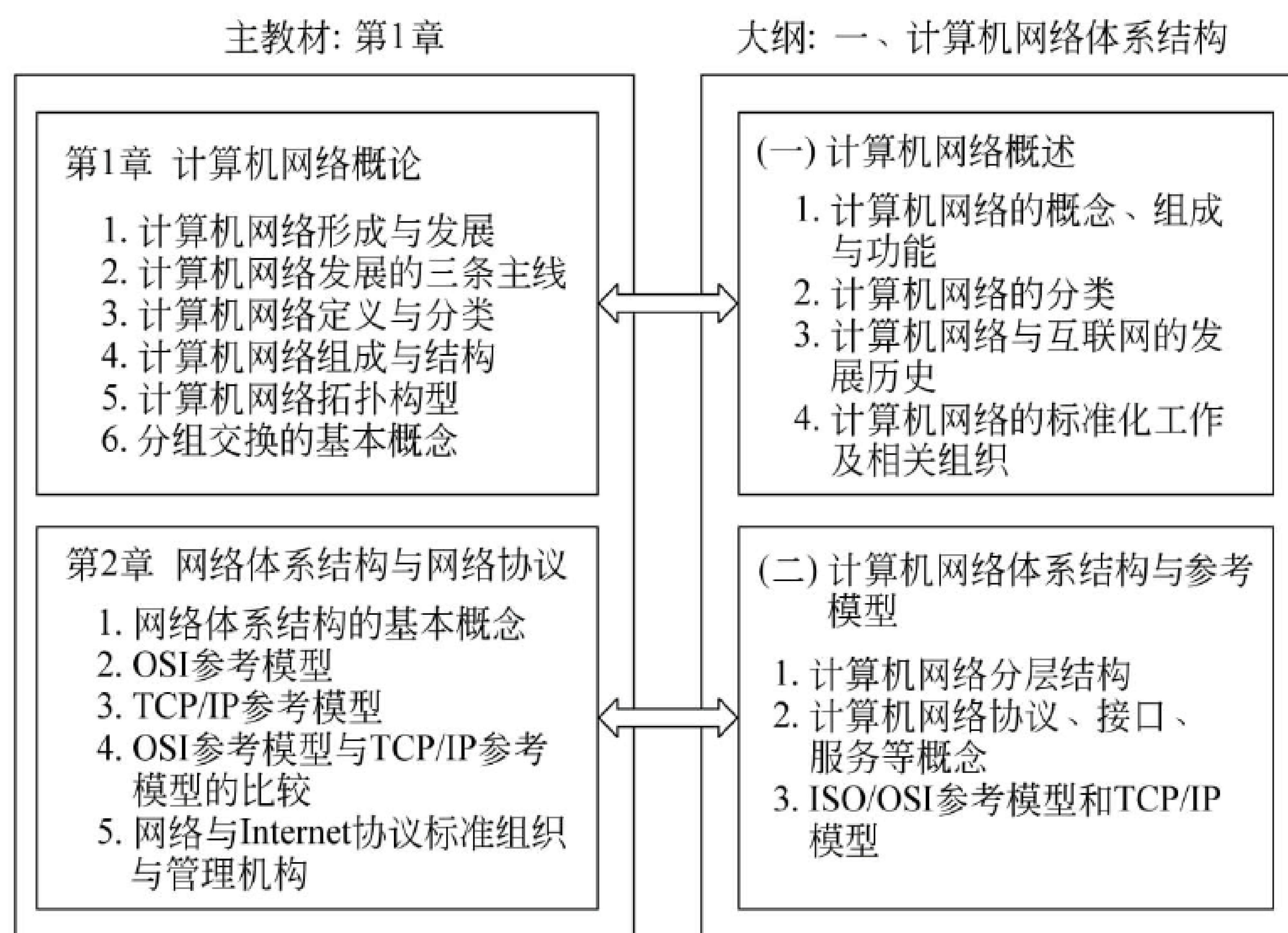


图 0-13 计算机网络体系结构与第 1 章之间的对应关系

概述中的前 3 个方面问题。第 4 个方面的问题在第 2 章有系统的讨论。

2. 第二个知识点：计算机网络体系结构与参考模型

计算机网络体系结构与参考模型又包括以下 3 个方面的内容：

(1) 计算机网络分层结构。

(2) 计算机网络协议、接口、服务等概念。

(3) ISO/OSI 参考模型和 TCP/IP 模型。

教材中第 2 章讨论了网络体系结构的基本概念，协议、接口与服务的基本概念，OSI 参考模型、TCP/IP 参考模型，以及网络协议标准化组织。因此，复习第 2 章能够覆盖大纲“计算机网络体系结构与参考模型”中涉及的所有的知识点。

0.8.3 物理层

大纲的第二部分是“物理层”。它又是由 3 个知识点组成。图 0-14 给出了物理层与教材的第 3 章之间内容上的对应关系。

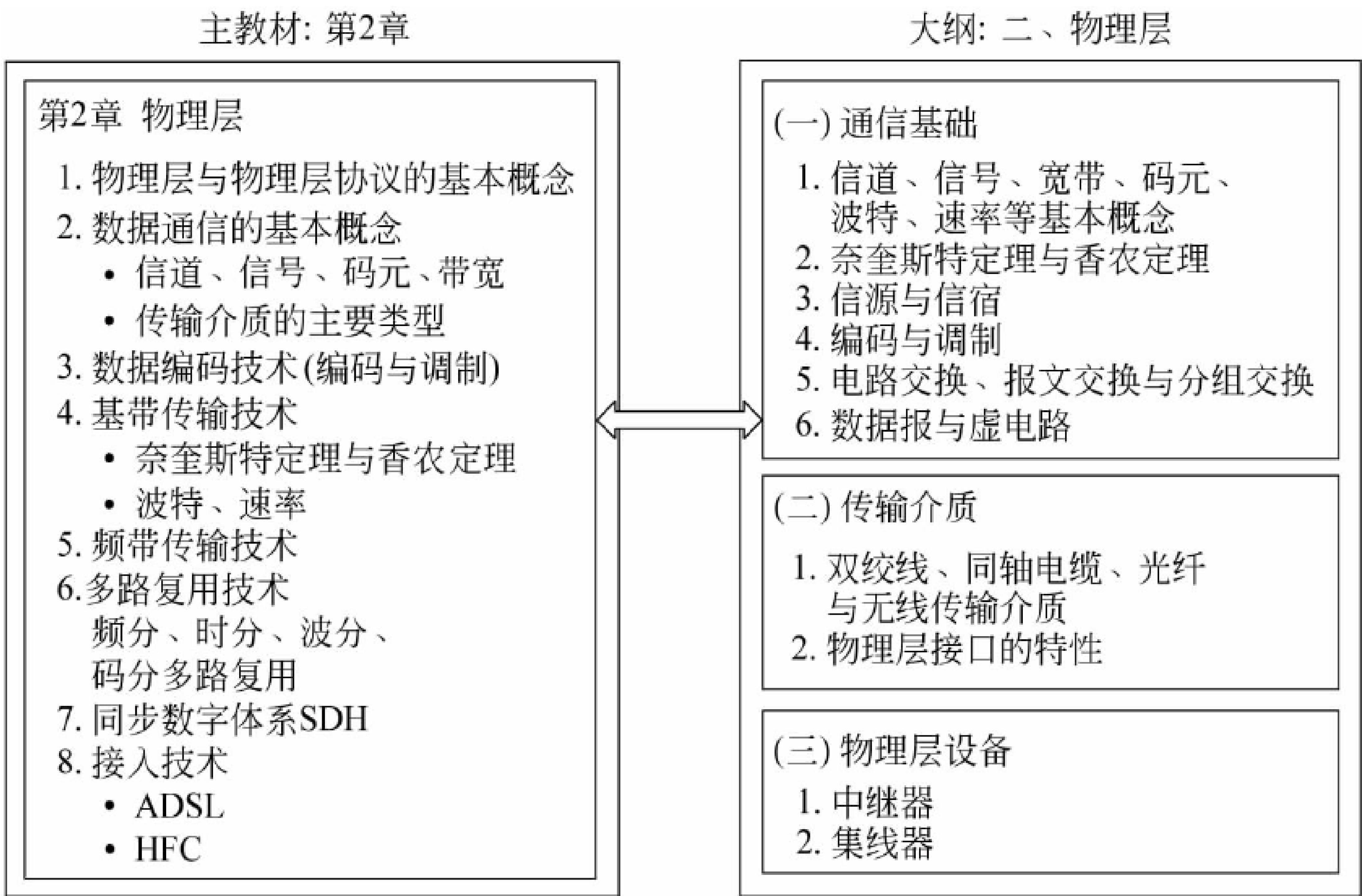


图 0-14 物理层与第 2 章内容的对应关系

1. 第一个知识点：通信基础

通信基础又包括以下六个方面的内容：

(1) 信道、信号、宽带、码元、波特、速率等基本概念。

(2) 奈奎斯特定理与香农定理。

(3) 信源与信宿。

(4) 编码与调制。

(5) 电路交换、报文交换与分组交换。

(6) 数据报与虚电路。

2. 第二个知识点：传输介质

传输介质又包括以下两个方面的内容：

(1) 双绞线、同轴电缆、光纤与无线传输介质。

(2) 物理层接口的特性。

3. 第三个知识点：物理层设备

物理层设备又包括以下两个方面的内容：

(1) 中继器。

(2) 集线器。

0.8.4 数据链路层

大纲的第三部分是"数据链路层"，它由8个知识点组成。图0-15给出了数据链路层与教材的第4、5章之间内容上的对应关系。大纲是将点-点链路与广播链路，以及广域网、局域网的数据链路层放在一个部分，对应一般的教材可能是1章，也可能是2章。对应本书选择的参考教材是对应于4、5章。



图0-15 数据链路层与第3、4章内容对应关系

1. 第一至四个知识点：数据链路层功能、组帧、差错控制、流量控制与可靠传输

前四个知识点包括数据链路层功能、组帧、差错控制、流量控制与可靠传输等内容。从技术分类的角度，它们是属于数据链路层基本概念，以及点-点链路与广域网数据链路层的协议的问题。这些内容在参考教材中集中在第4章中讨论。

2. 第五、六个知识点：介质访问控制、局域网

第五、六个知识点包括介质访问控制、局域网。从技术内涵讲，介质访问控制是解决局域网多结点共享和争用的控制算法与协议，讨论局域网原理时必然会涉及的理论基础问题。在参考教材中集中在第5章中进行了系统的讨论。

3. 第七个知识点：广域网

第七个知识点广域网包括广域网基本概念、PPP协议、HDLC协议与ATM网络基本原理。这一部分内容在参考教材中分别在第1章讨论了广域网的基本概念、ATM基本工作原理，第4章讨论了HDLC协议与PPP协议。

4. 第八个知识点：数据链路层设备

第八个知识点数据链路层设备包括网桥与交换机两个部分。网桥部分注明了网桥的基本概念、透明网桥与源选径网桥，以及对应的生成树算法与源选径算法。这些内容在参考教材的第4章中均有系统的讨论。

0.8.5 网络层

大纲的第四部分是“网络层”，它由8个知识点组成。图0-16给出了网络层与教材的第5章之间内容上的对应关系。

1. 第一个知识点：网络层的功能

网络层的功能主要包括异构网络互联、路由与转发、拥塞控制等内容。

2. 第二个知识点：路由算法

路由算法主要包括静态路由与动态路由、距离-向量路由算法、链路状态路由算法、层次路由等内容。这些内容与第五部分路由选择协议表面上看有重叠的部分。但是从大纲整体安排角度，这一部分更多体现在路由选择算法的概念上，为第五部分的路由选择算法作为基础。

3. 第三个知识点：IPv4

IPv4部分主要包括IPv4分组、IPv4地址与NAT、子网划分与子网掩码、CIDR、ARP协议、ICMP协议等内容。

4. 第四个知识点：IPv6

IPv6主要包括IPv6的主要特点、IPv6地址等内容。

5. 第五个知识点：路由选择协议

路由选择主要包括自治系统、域内路由与域间路由、RIP路由协议、OSPF路由协议、BGP路由协议等内容。

6. 第六个知识点：IP组播

IP组播主要包括组播的概念、组播路由算法等内容。

7. 第七个知识点：移动IP

移动IP主要包括移动IP的概念、移动IP通信过程等内容。

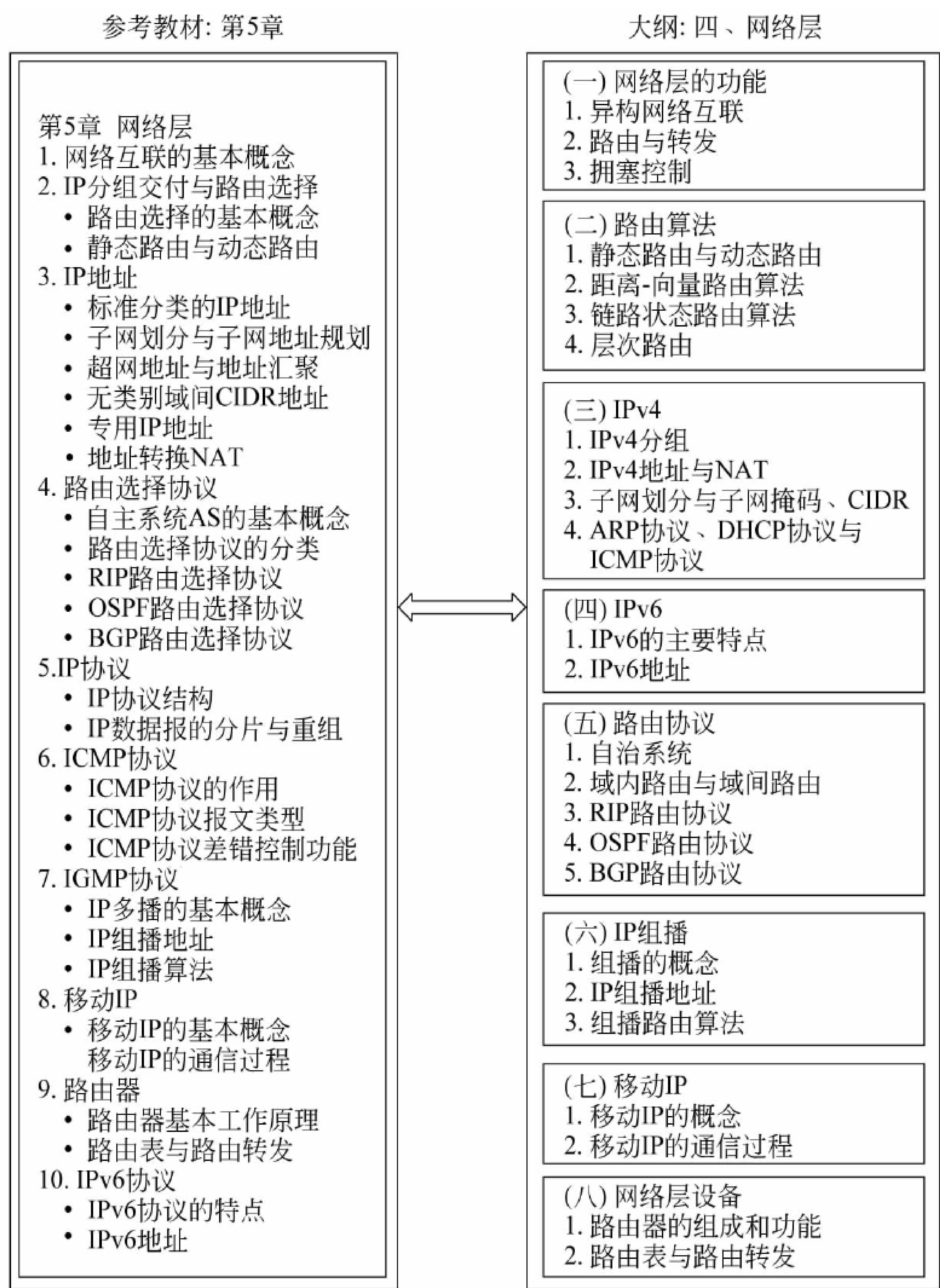


图 0-16 网络层与第5章内容对应关系

8. 第八个知识点：网络层设备

网络层设备主要包括路由器的组成与功能、路由表与路由转发等内容。

参考教材的第5章基本上都涵盖了以上内容。

0.8.6 传输层

大纲的第五部分是“传输层”，它由3个知识点组成。图0-17给出了传输层与第6章内容对应关系。

1. 第一个知识点：传输层提供的服务

传输层提供的服务主要包括传输层的功能、传输层寻址与端口、无连接服务与面向连接服务等内容。

2. 第二个知识点：UDP 协议

UDP协议主要包括UDP数据报、UDP校验等内容。

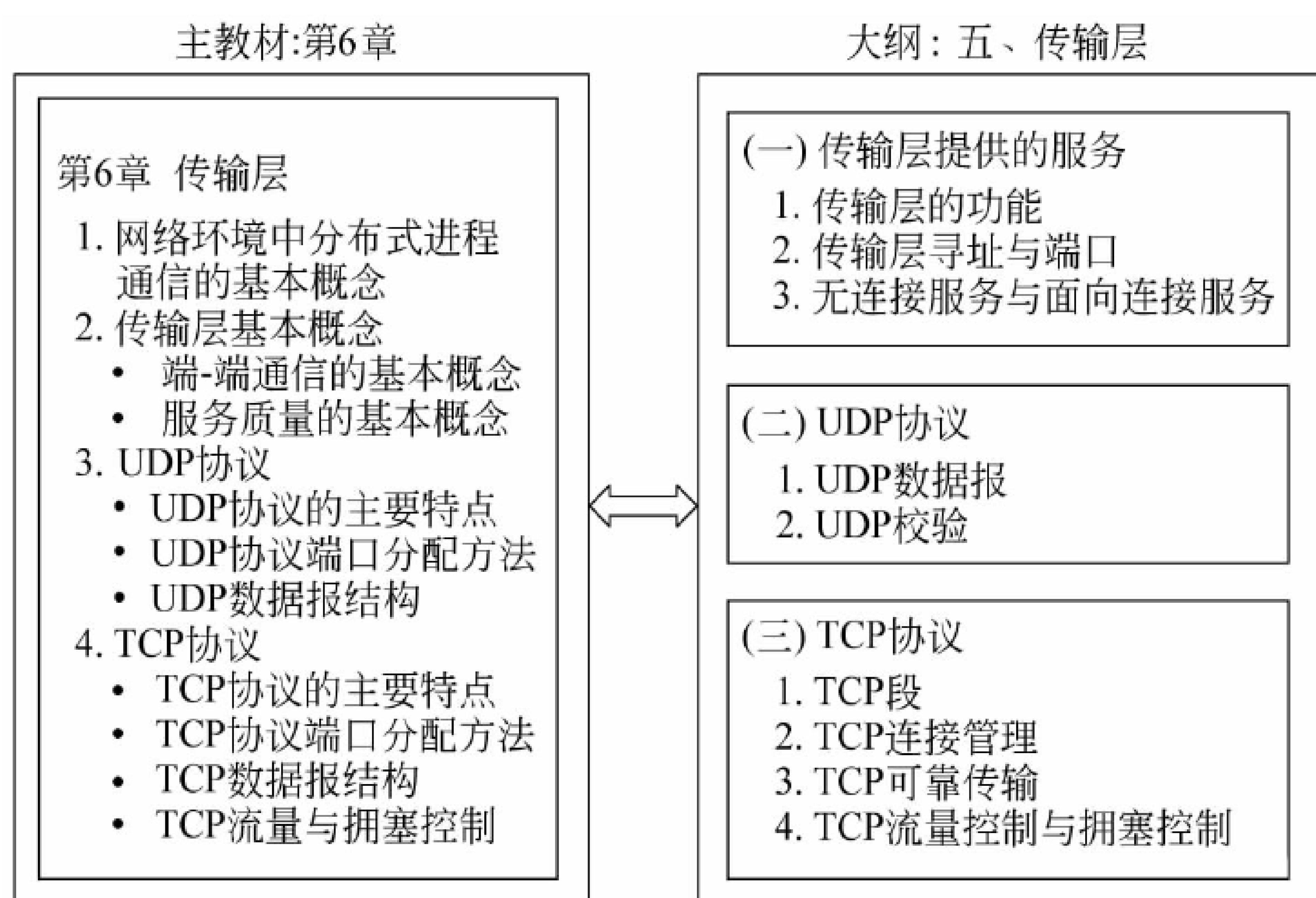


图 0-17　传输层与第 6 章内容对应关系

3. 第三个知识点：TCP 协议

TCP 协议主要包括 TCP 连接管理、TCP 可靠传输、TCP 流量控制与拥塞控制等内容。参考教材的第 6 章基本上都涵盖了以上内容。

0.8.7　应用层

大纲的第六部分是“应用层”，它由 5 个知识点组成。图 0-18 给出了应用层与第 7 章内容对应关系。

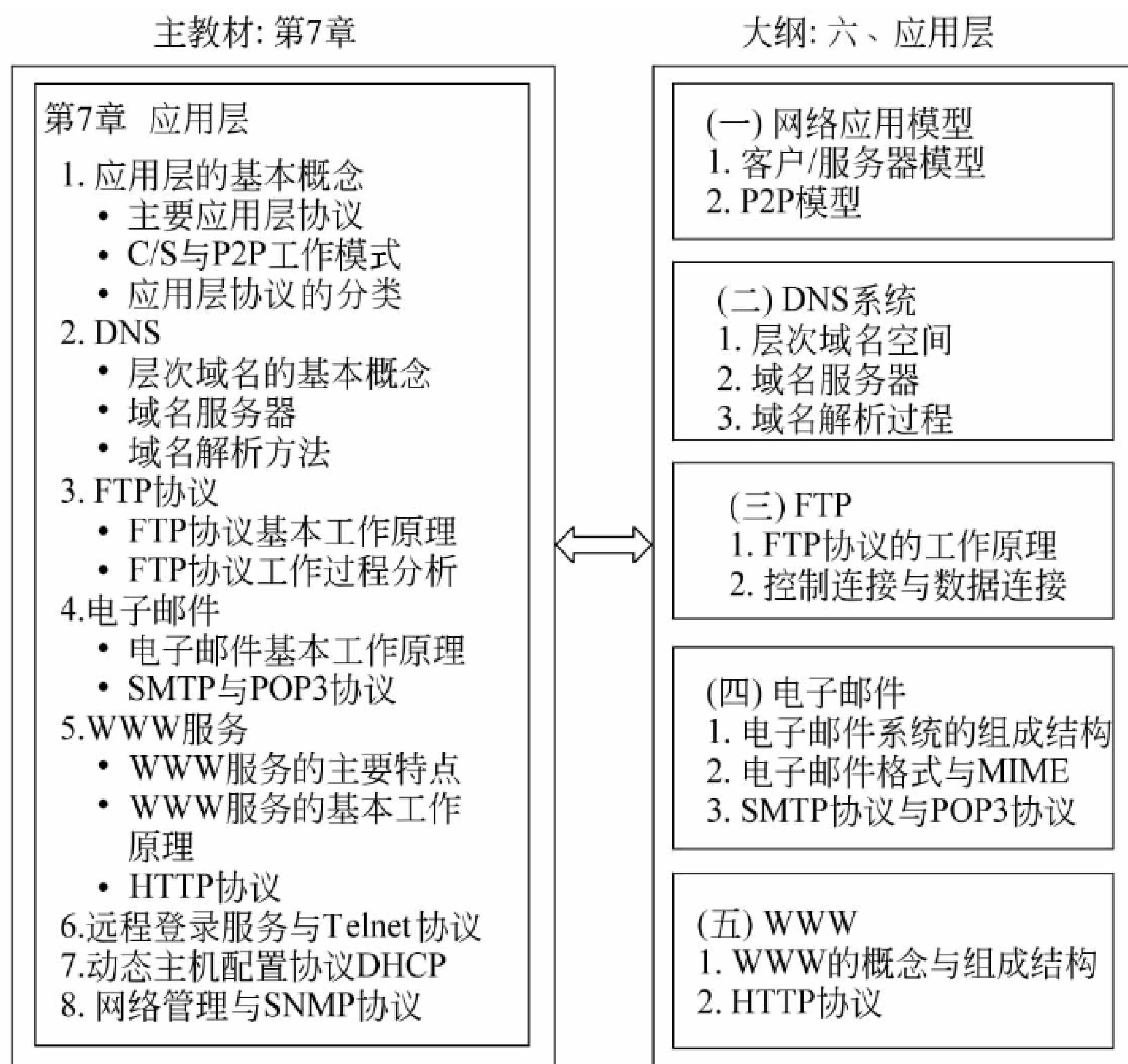


图 0-18　应用层与第 7 章内容对应关系

1. 第一个知识点:网络应用模型

网络应用模型主要包括客户/服务器模型、P2P模型等内容。

2. 第二个知识点:DNS系统

DNS系统主要包括层次域名空间、DNS服务器、域名解析过程等内容。

3. 第三个知识点:FTP

FTP主要包括FTP协议的工作原理、控制连接与数据连接等内容。

4. 第四个知识点:电子邮件

电子邮件主要包括电子邮件系统的组成结构、电子邮件格式与MIME协议等内容。

5. 第五个知识点:WWW

WWW主要包括WWW的概念与组成结构、HTTP协议等内容。

参考教材的第7章基本涵盖了以上内容。

"基本涵盖"是指教材对以上内容都做了较为系统地讨论,但是不同教材的侧重点、对问题讨论的深度差异都很大,因此在本书中将通过例题与习题对基本概念、原理与方法,以及一些综合应用进行强化训练之外,对需要强调的内容做一些补充。同时,为了适应考纲的要求,本书也将对相关的问题作适度的扩展,补充了个别的例题与解析,这些在书中均做出了说明。复习的过程应该不单单是为了应试,同时也应该是增进对知识与技能理解、掌握的过程。

0.8.8 对于复习、备考的建议

对于复习、备考,读者需要注意以下几个问题:

(1) 由于"计算机科学与技术专业基础综合考试大纲"包括4门课程,学生在准备考试之前,首先需要了解大纲对一门课程考查的要求,然后再仔细研读考查范围与具体的知识点结构,结合自己在读这门课程时对各个知识点的掌握情况,巩固已经掌握的比较好的知识点,找出自己认为欠缺的部分,突出重点,集中精力,通过复习,力求基本达到考纲的基本要求。四门核心课程,如此之多的内容,如果没有重点,全面出击,不可能取得预期的结果。

(2) 不同的同学的一门课程学习和掌握的情况差异很大,应该有所区别。对于很多在网络知识上掌握得比较好的同学,他们可以通过快速地浏览例题和完成练习题,复习知识,找出不足,有重点的提高。对于网络知识上掌握得一般的同学,他们可能心里没底,不知道如何复习。建议这一类同学跟着辅导教材,不要先看例题解析,而是自己先做,然后与例题解析的过程去比较,在比较过程中发现自己知识点的缺陷,即时补充。然后通过独立完成练习题去检查自己的掌握情况,进一步找出不足,第二遍有重点的复习,争取能够在不太长的时间取得比较好的进步。

(3) 复习、备考的过程是一个检验自己对重要知识的真实掌握情况的考查、补充和完善的过程。复习、备考最忌讳只看教科书,不做练习。因为入学考试的试题一定是对同学对这门课程重要概念、基本理论、技能的综合考查。例如2009年入学考试命题可以看出,除去考查一部分重要的概念之外,重点是考查学生对于网络技术综合应用的能力。因此,复习、备考的过程一定要通过在完成例题和习题的过程中发现自己知识上的缺陷和综合应用的能力。例题中有大量综合应用的题目,练习中又对重点和容易混淆的问题加重训练。读者如果在完成例题与习题的过程中发现还需要补充知识,可以结合课本进一步复习。大纲公布

的试题示例中表示试题类型只有两种：单选题与综合应用题。例题与练习题的命题已经参考了大纲公布的命题格式，这样在完成例题与练习的过程也可以进一步地适用考试。因此希望读者能够在对于例题和做习题一样，先自己完成，做不出来再看解析，找出自己在哪些知识上需要加强，哪些错误的概念需要纠正；做出来的可以与解析比较，看看哪种方法更好。在完成例题与习题的过程中，再结合课本进行复习，这样会起到事半功倍的效果。

第 1 章 计算机网络概论

第一部分　学习目的、要求与知识点结构

1. 学习目的

根据本书整体结构的安排，第 1 章回答了两个最基本的问题：什么是计算机网络？什么是网络协议与网络体系结构？

第 1 章的学习目的是：了解网络形成与发展历史，掌握网络定义、分类与拓扑构型等几个基本知识，对现代计算机网络结构有一个清晰的认识，对计算机网络核心概念——分组交换概念与技术有一个深入的理解，帮助读者对计算机网络技术建立一个全面与正确的认识。

掌握网络技术中两个最基本的概念：网络体系结构与网络协议，以便读者对计算机网络的工作原理和实现技术建立一个整体的概念，为以后的学习打下基础。

2. 学习要求

（1）了解：计算机网络的形成与发展过程。

（2）掌握：计算机网络的定义与分类。

（3）掌握：计算机网络的组成与结构的基本概念。

（4）掌握：计算机网络拓扑构型的定义、分类与特点。

（5）掌握：分组交换的基本概念。

（6）掌握：网络协议与网络体系结构的基本概念。

通过本章的学习，帮助学生掌握网络的基本概念，启发学习兴趣，为进一步深入地学习打下基础。

3. 本章知识点的组织与结构

在组织本章内容时，作者注意了以下几个问题。

（1）计算机网络和 Internet 是一个对社会的发展有着深远的影响的技术，它对信息技术与产业的发展有着不可估量的作用，因此我们需要从一个更大的视野去认识网络技术。

（2）本章从宏观的层面介绍了计算机网络和 Internet 技术发展与演变的历史过程、三条技术发展主线，从微观的层面讨论了计算机网络的定义、结构、分类与拓扑等基本概念。

（3）本章的重点是要求学生掌握计算机网络的定义、结构、分类与拓扑、网络协议与网络体系结构等基本概念。

第一个问题："什么是计算机网络"，对应的知识点结构如图 1-1 所示。

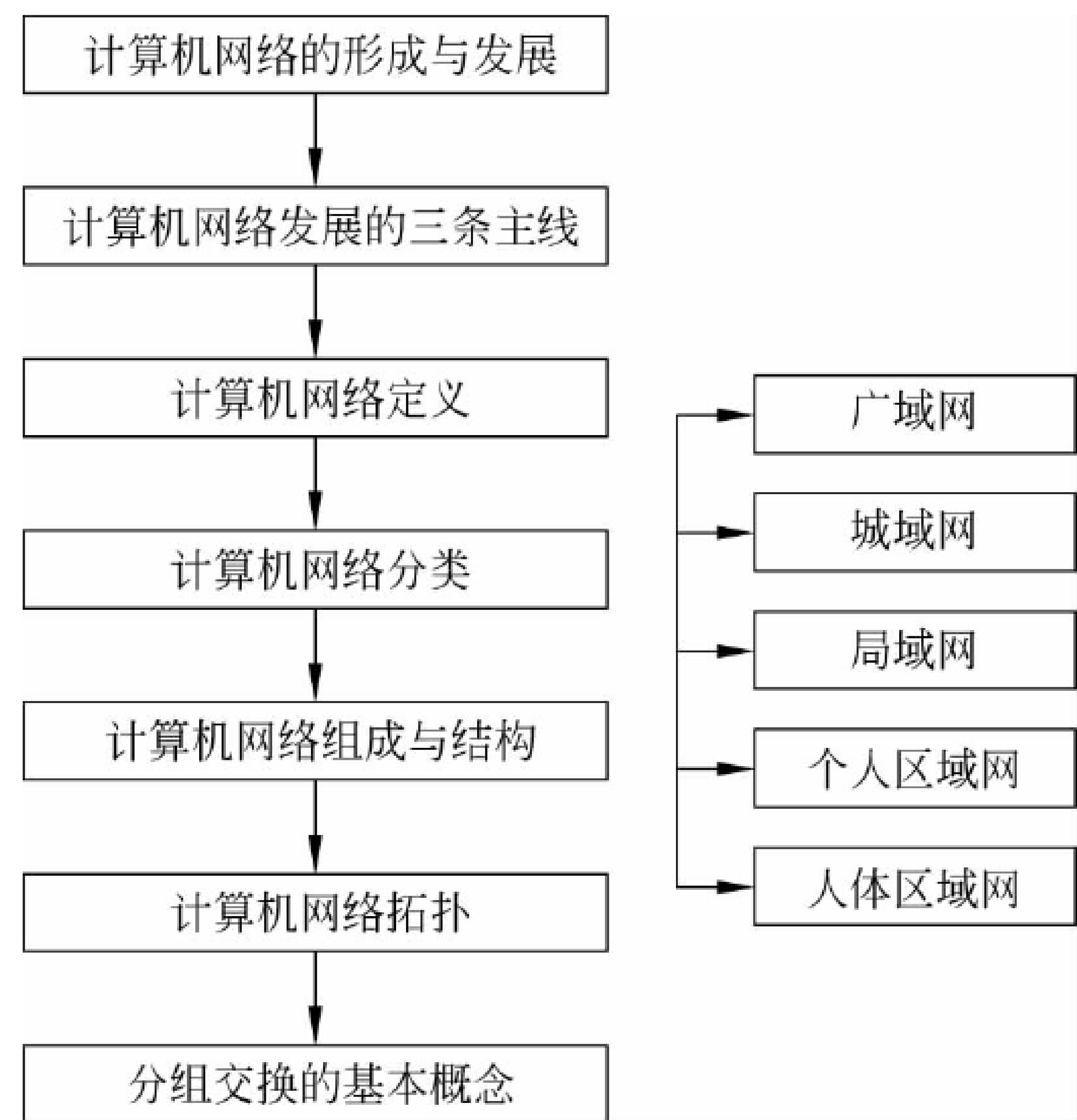


图 1-1　问题 1 对应的知识点结构

第二个问题："什么是网络协议与网络体系结构"，对应的知识点结构如图 1-2 所示。

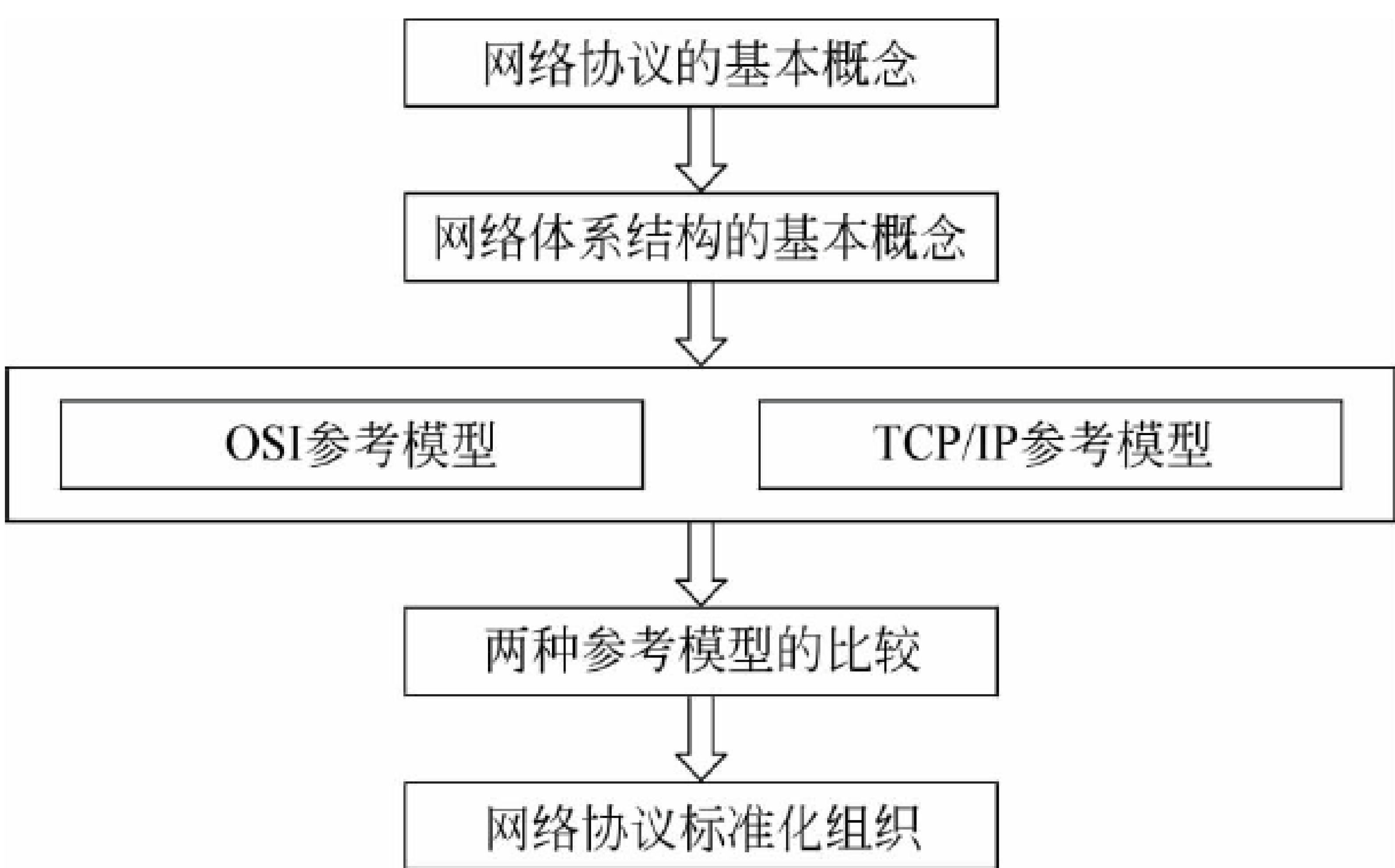


图 1-2　问题 2 对应的知识点结构

第二部分　教学内容问答

问题 1-1：分组交换技术经历了怎样的发展与演变过程？

在讨论分组交换技术与计算机网络发展的关系时，学术界普遍的认识是：分组交换奠定了计算机网络发展的基础，标志着现代通信网络技术时代的诞生。这种评价是恰如其分的。因为无论现代网络技术如何变化，分组交换的概念没有变化，计算机网络、电信网络与电视传输网的数据传输都是建立在分组交换技术的基础上的，这是计算机网络研究对现代通信技术发展的重大贡献。教师在讲授分组交换技术时，需要深入了解一些分组交换技术的背景知识，研究出现的背景，研究的过程，以及在这个过程中做出贡献的重要科学家。

1. 分组交换网研究的背景

世界上第一台电子数字计算机 ENIAC 出现在 1946 年,但是通信技术的发展要比计算机技术早很长时间。在很长的一段时间中,这两种技术之间并没有直接联系,处于各自独立发展的阶段。当计算机与通信技术都发展到一定程度时,并且社会上出现了新的应用需求时,人们就会产生将两项技术交叉融合的想法。计算机网络就是计算机技术与通信技术高度发展、深度融合的产物。

20 世纪 50 年代初,由于美国军方的需要,美国半自动地面防空(Semi-Automatic Ground Environment,SAGE)系统将远程雷达信号、机场与防空部队的信息,通过总长度为 241 万千米的通信线路(包括有线与无线通信),传送到位于美国本土的一台 IBM 计算机上来处理,这项研究开始了计算机技术与通信技术结合的尝试。随着 SAGE 系统的实现,美国军方又考虑将分布在不同地理位置的多台计算机通过通信线路连接成计算机网络的需求。

20 世纪 60 年代中期,全世界正处于“冷战”高潮时期。1957 年 10 月,苏联在拜科努尔航天中心向太空成功发射了第一颗人造卫星 Sputnik(史伯尼克),美国朝野为之震惊,事关国家安全危机的阴云笼罩着整个美国。他们的第一反应是成立一个专门的国防研究机构,即美国国防部高级研究计划署(Advanced Research Projects Agency,ARPA),办公地点设在五角大楼内。由于它是美国国防部的一个机构,因此也有文献使用 DARPA 表示。总之,ARPA 是一个科研管理机构,它没有实验室与科学家,只是通过签订合同和发放许可的方式,选择一些大学、研究机构和公司为该机构服务。

在与苏联的军事力量竞争中,美国军方发现需要一个专门用于传输军事命令与控制信息的网络。他们希望这种网络在遭到核战争或自然灾害后,在部分通信设备或通信线路遭到破坏的情况下,通信网络系统仍然能利用剩余的网络设备与通信线路继续工作,这个网络也被称为“网络可生存(Network Suvivality)”系统。这种要求是传统的通信线路与电话交换网所无法实现的。当时美国军方的通信主要依靠电话交换网,但是电话交换网是相当脆弱的。由于电话交换网是以每个电话交换局为中心组成的星—星结构,从而形成一个覆盖全国的层次型结构的电话通信系统,因此电话交换网的一个中继线路或交换机的损坏,尤其是几个关键长途电话局遭到破坏,就有可能导致整个电话通信的中断。针对这种情况,美国国防部开始着手进行新的通信网络技术——分组交换网的研究工作。

2. 分组交换网概念的提出

分组交换概念的研究工作主要集中在 1961—1967 年。ARPANET 的实现证明了分组交换理论的正确性。

在讨论 ARPANET 的产生背景时,一定会涉及兰德(Research and Development,RAND)公司与 J. C. R. Licklider、Paul Baran、Donald Davies、Wesley Clark、Larry Roberts 等几位科学家。实际上,分组交换概念与技术的研究是由三部分科学家并行开展的。

1) MIT 与 J. C. R. Licklider、Lawrence Roberts 的研究工作

MIT 在计算机网络与分组交换技术早期的研究工作主要集中在 1961—1967 年。贡献比较突出的两位科学家是 J. C. R. Licklider(约瑟夫·利克莱德)与 Lawrence Roberts(劳伦斯·罗伯茨)。

20 世纪 50 年代,J. C. R. Licklider 作为 MIT 计算机系的副教授参加了半自动地面防空

(SAGE)系统的研究工作(如图 1-3 所示)。这对于他在 1962 年提出“星际计算机网络(Intergalactic Computer Network)”的概念是有很大帮助的。在有关“星际计算机网络”的备忘录中,他描述了在全球范围内将很多计算机互连起来,使得每个人从任何地点很快能够得到数据与程序的设想。ARPA 的核心机构之一是信息处理办公室(Information Processing Techniques Office,IPTO),它主要负责计算机网络、超级计算机等研究课题。1962 年,J. C. R. Licklide 离开 MIT 加入 ARPA,并在后来成为 IPTO 的首席执行官。他关于计算机网络的设想正好与 ARPA 对“可生存系统”研究的需求是一致的。他的“星际计算机网络”思想对后来的 ARPANET 研究产生了重要的影响。

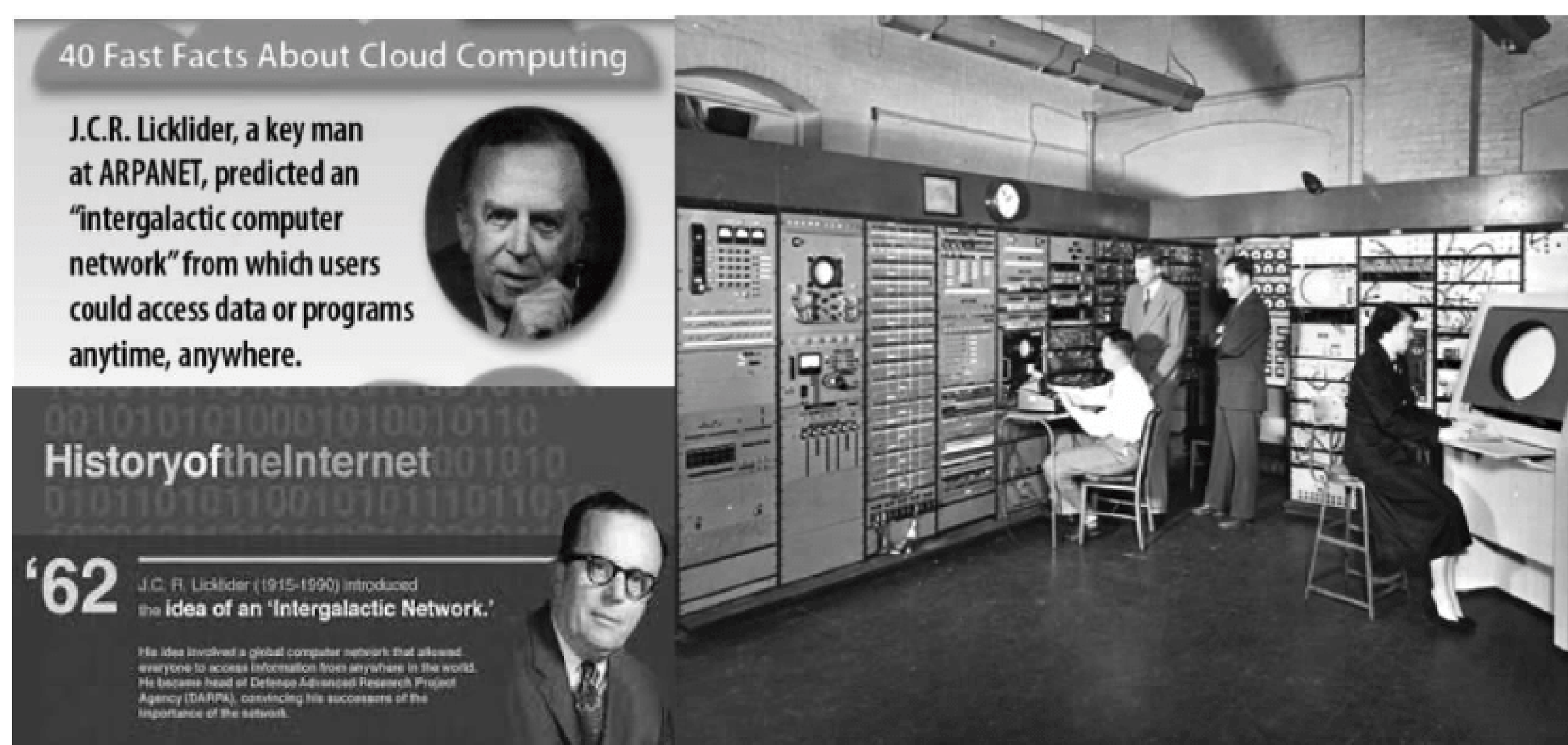


图 1-3 J. C. R. Licklider 与 MIT 的 SAGE 研究室

另一位在计算机网络与分组交换技术研究方面做出重要贡献的科学家是 Lawrence Roberts。1963 年,Lawrence Roberts 获得 MIT 博士学位之后加入了 MIT 的林肯实验室。他在 1961 年发表了题为“*Information Flow in Large Communication Network*”的论文,全面阐述了关于分组交换的思想。1965 年,Lawrence Roberts 与 Thomas Merrill 在分组交换技术上迈出了划时代的一步。他们将位于 MIT 与加州的两台分时计算机通过电话线路连接起来,存取数据、运行程序。他们的研究工作第一次向世人证实了分组交换理论的可信性。当时他们就将相互间传送文件的约定称为“协议(Protocol)”。

作为 ARPA 的首席科学家,1967 年 10 月 Lawrence Roberts 提交了第一个分组交换网 ARPANET 研究计划,并组织完成了 ARPANET 结构与通信协议的研究规划、设计工作。图 1-4 是在 ARPANET 的研究工作中做出重要贡献的计算机科学家 Lawrence Roberts 的照片。

2) RAND 公司与 Paul Baran 的研究工作

1948 年 5 月,RAND 公司在 Santa Monica 成立,它是美国政府在第二次世界大战后成立的一个战略研究机构。当时 RAND 公司的研究工作重点是冷战时期的军事战略问题。1960 年,美国国防部授权 RAND 公司寻找一种有效的通信网络解决方案。

当时 Paul Baran(保罗·巴兰)在 RAND 公司的数学学部的计算机科学分部工作。他最有兴趣的研究课题是在受到核攻击之后,如何维持网络的通信能力。1962 年,Paul Baran 为 RAND 公司写了 11 份报告,讨论了“包交换(Packet Switching)”与“存储转发(Store and Forward)”的基本工作原理。其中影响最大的是 1964 年 3 月发表的“论分布式通信网络

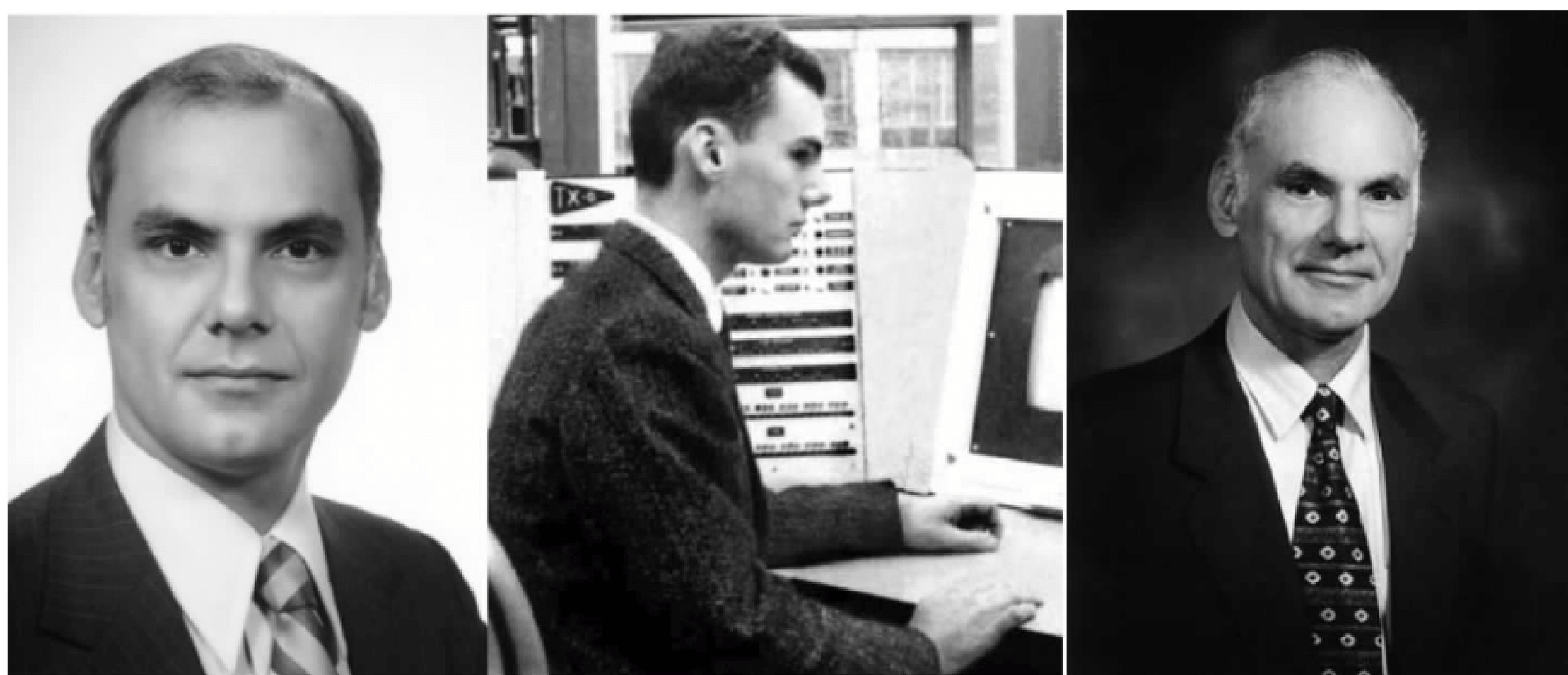


图 1-4 在 ARPANET 研究工作中做出重要贡献的 Lawrence Roberts

(*On Distributed Communication Networks*)”。Paul Baran 利用计算机和容错技术设计了比电话交换网更为健壮的通信网络。不过,当时他的许多同事对计算机知识了解甚少,因此多数人认为他的研究成果没有意义。尽管如此,Paul Baran 还是继续研究这个问题,撰写学术论文,并进一步完善自己的设计思想。图 1-5 是对分组交换网理论做出重要贡献的 Paul Baran。

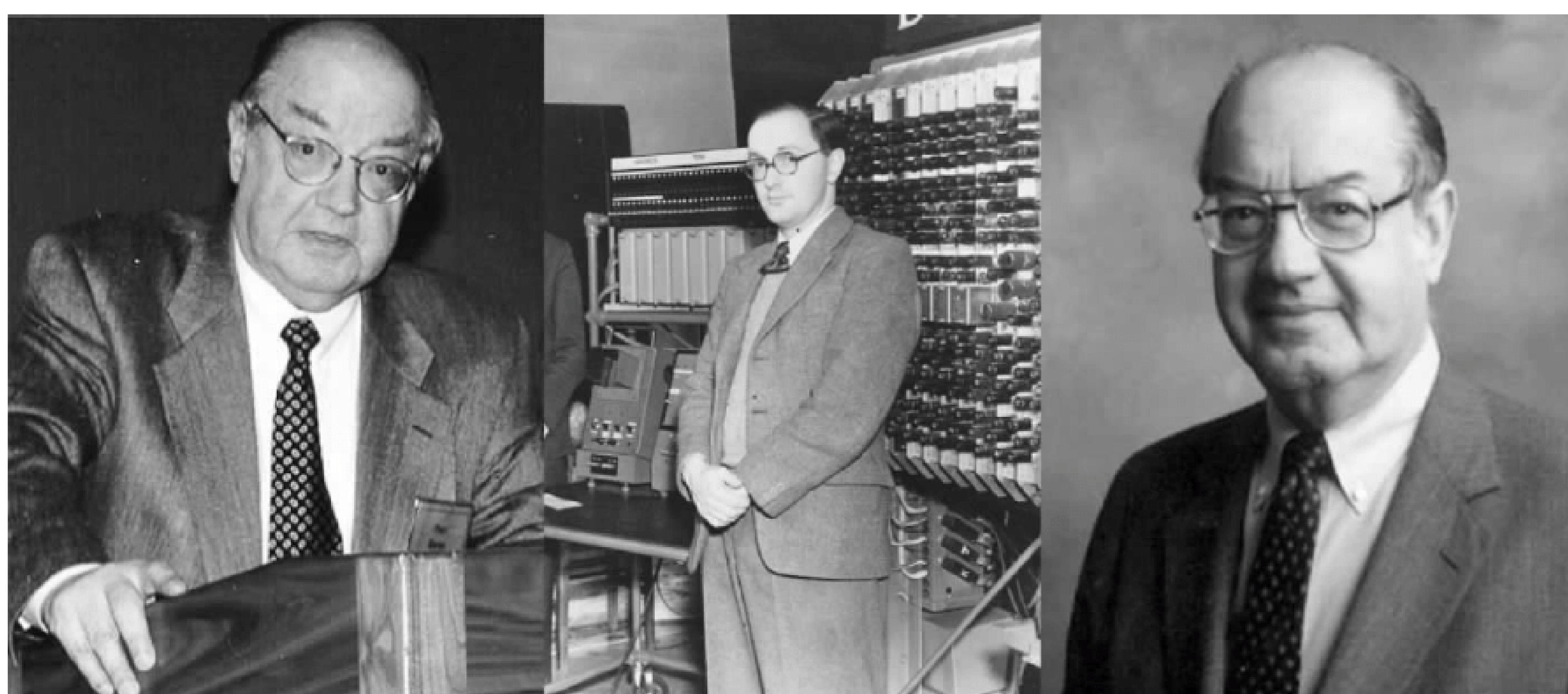


图 1-5 对分组交换网理论做出重大贡献的 Paul Baran

3) NPL 与 Donald Davies 的研究工作

在 1967 年下半年的 ACM SIGOPS 会议上,英国国家物理实验室(National Physical Laboratory,NPL)的 Donald Davies(唐纳德·戴维斯)的文章描述了一个类似的网络系统。他引用了 Paul Baran 的研究成果,并且在 NPL 建立了一个实验系统。这个系统证实了分组交换思想的正确性。图 1-6 是对分组交换网理论做出重大贡献的计算机科学家 Donald Davies。

3. 分组交换网技术的研究

当时,构建通信网络有两种基本的拓扑构型模式:集中式(Centralized)和非集中式(Decentralized)。图 1-7 给出了集中式和非集中式的拓扑构型。在集中式网络中,所有节点都与中央交换节点相连,所有数据都要发送给中央节点,再通过它传送到目的节点。如果中心节点受到损坏或功能不正常,所有通信就会完全中断。非集中式网络使用了若干个中心

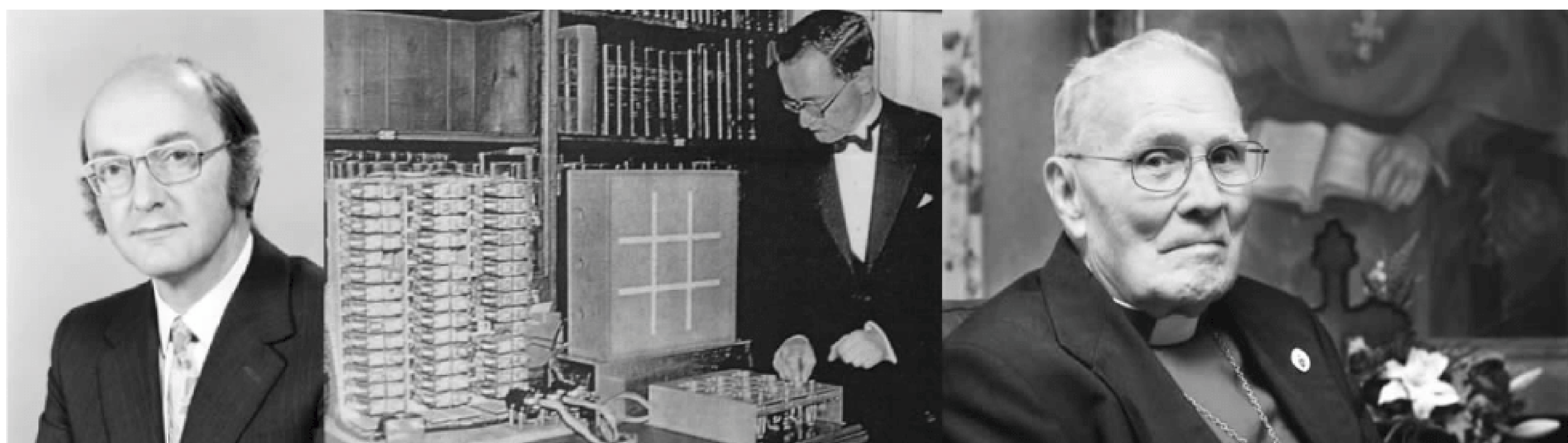


图 1-6　对分组交换网理论做出重大贡献的 Donald Davies

节点，相当于许多集中式网络连接起来，固有的缺点仍然无法避免。当时美国的通信系统基本上是非集中式网络模式。

Paul Baran 提出了第三种设计方案——分布式网络(Distributed Network)结构。他的基本设计思想是：分布式网络没有中心交换节点，网络中每个节点都有若干个邻节点，从而形成网状结构。通信网络必须让几百个主要节点之间都能通信。每个节点可以通过多条路由去发送数据。这样，当网络中某一个或几个节点损坏时，还有其他路由可以使用。图 1-8 给出了分布式网络的拓扑构型。显然，这是一种高度分布和容错的网状结构设计方案。

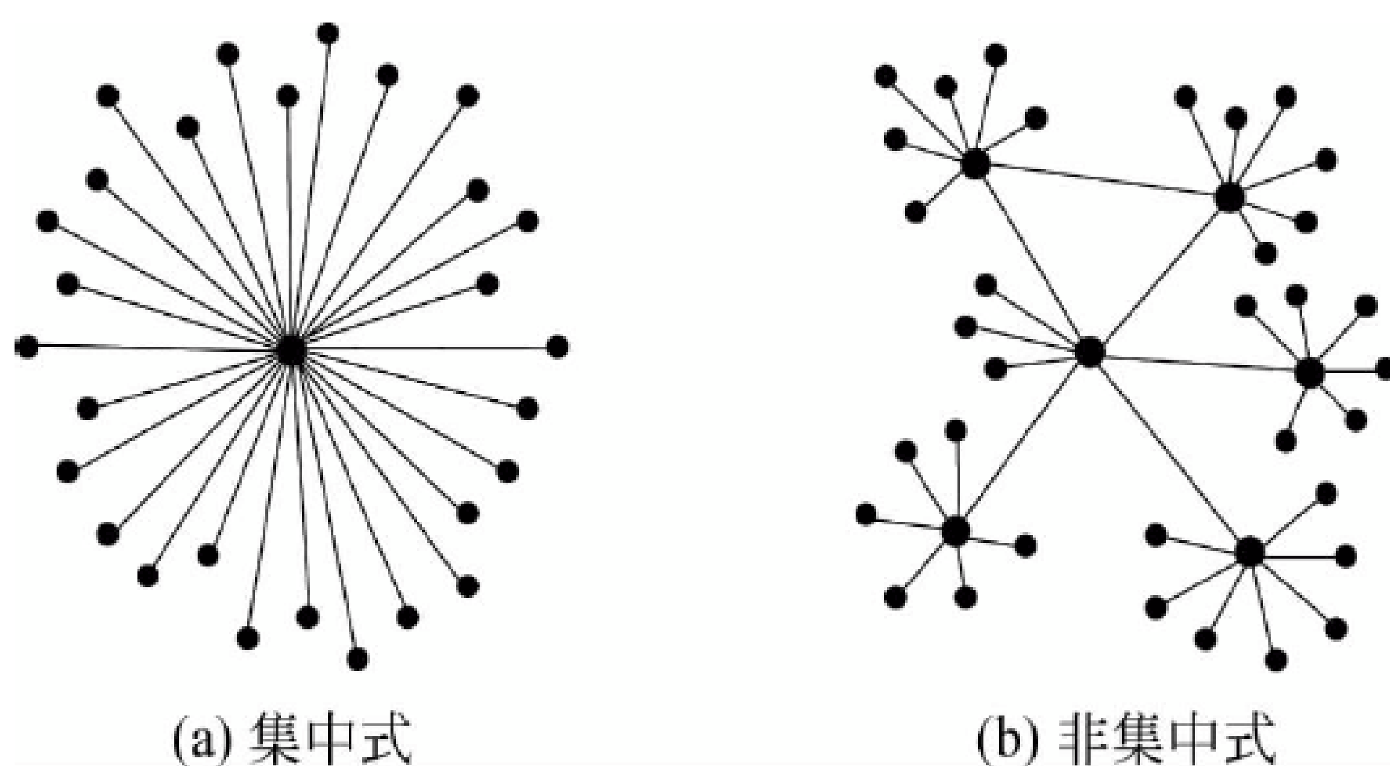


图 1-7　集中式和非集中式的拓扑构型

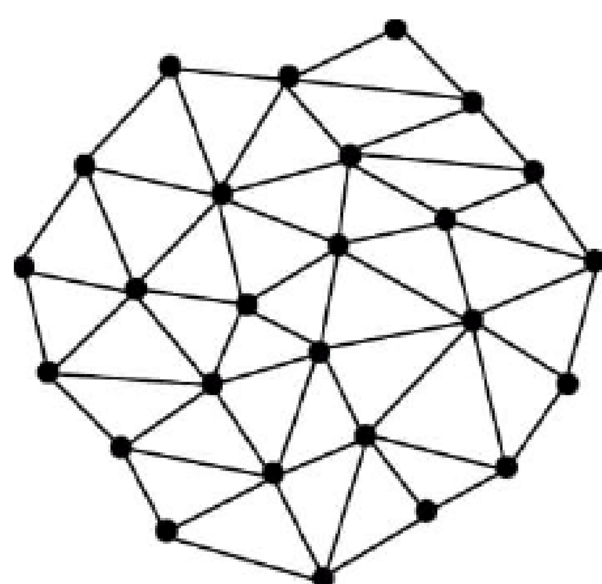


图 1-8　分布式网络的拓扑构型

Paul Baran 建议在新的通信系统中采用数字的分组交换技术。他想象的是一个没有集中节点的网络，网络中节点之间传输的数据按事先规定的格式封装在一种叫作“分组(Packet)”或“包”的传输单元中。分组中包含源地址与目的地址。传输网络中的节点使用路由选择算法(Routing Algorithms)为每一个分组选择一条到达目的节点的“最佳”路径。每一个传输路径中的节点按照存储转发方法将分组发送到下一个节点。如果传输路径中有一个节点或线路损坏，分组可以通过一种“动态路由(Dynamic Routing)”算法来修改分组的路由，绕道而行，最终到达目的节点。

Paul Baran 向美国国防部的报告中详细地阐述了他的建议。美国国防部接受了 Paul Baran 的建议，并请 AT&T 公司与国家电话局建立一个原型系统。AT&T 公司最终没有采用 Paul Baran 的方案。但是，Paul Baran 提出的分组交换网的设计思想为计算机网络的研究指出了正确的方向。

在讨论这一段研究工作经历的时候，我们需要注意这样一个细节。Paul Baran 于 1926 年出生于波兰，1959 年获 UCLA 电气工程硕士学位。Paul Baran 在进入 RAND 公司之前，曾经在美国 Eckert-Mauchly Computer 公司工作，参与早期 UNIVAC 计算机的研制

工作,有着很好的计算机技术背景。Donald Davies 出生于英国,在伦敦帝国学院获数学、物理双学位。1947年进入NPL,在图灵的领导下研制英国第一台计算机 Pilot ACE。从 Paul Baran 与 Donald Davies 这两位对分组交换理论发展起到重要作用科学家的教育经历中,可以清楚地认识到:正是由于他们有很好的计算机知识背景,又能够将计算机与通信这两个相互独立发展学科的知识有机地交叉与融合起来,他们才有可能在推动计算机网络理论与实现技术的领域中取得举世瞩目的研究成果。

问题1-2:ARPANET 的研究经历了怎样的发展与演变过程?

ARPANET 的研究证实了分组交换概念的正确性,是第一个计算机网络的原型系统,为互联网的发展奠定了坚实的基础。研究这个问题需要注意以下几点。

1. ARPANET 基本设计思想

在讨论 ARPANET 研发过程时,一定会涉及另外两位计算机科学家——Robers Taylor(罗伯特·泰勒)与 Robert Kahn(罗伯特·卡恩)。

1966年,Robers Taylor 出任 ARPA 的信息处理办公室 IPTO 的第三任主任。当时五角大楼内的 IPTO 办公室就有分别连接 MIT、加州大学 Berkeley 分校和终端系统开发公司分时计算机系统的三种相互不能够兼容的终端设备。实现计算机的互连,为用户提供信息共享是这一代计算机科学家共同的心愿,Robers Taylor 决定迈出实现理想的第一步。

他向 ARPA 署长 C. Herzfeld(查尔斯·赫兹菲尔德)提出由 ARPA 出面建设一个小型的实验网络。C. Herzfeld 批准了 Robers Taylor 的建议,标志着 ARPA 正式立项支持 ARPANET(图1-9是 C. Herzfeld、Robers Taylor 与 ARPA 实验室的照片)。



图1-9 Robers Taylor、C. Herzfeld 与 ARPA 实验室

作为 IPTO 主任的 Robers Taylor 决定将林肯实验室的 Lawrence Roberts 调到 IPTO 办公室,来负责这项研究。Lawrence Roberts 当时正在林肯实验室研究两台异构计算机系统的互连问题。尽管最初 Lawrence Roberts 并不愿意调到 IPTO,但是当 Lawrence Roberts 来到五角大楼 IPTO 办公室之后,他马上就用秒表对五角大楼内部所有走廊进行了

测量,计算出各个办公室之间行走的最短路线(即“拉里路线”),为不同办公室计算机之间的通信寻找“最短路径”。

1967 年,Lawrence Roberts 提交了 ARPANET 最早的设计方案。应该说,ARPANET 的设计方案是在前期分组交换技术理论研究基础上形成的,它凝聚着一代计算机科学家的集体智慧。图 1-10 给出了 ARPANET 分组交换网的结构与原理示意图。

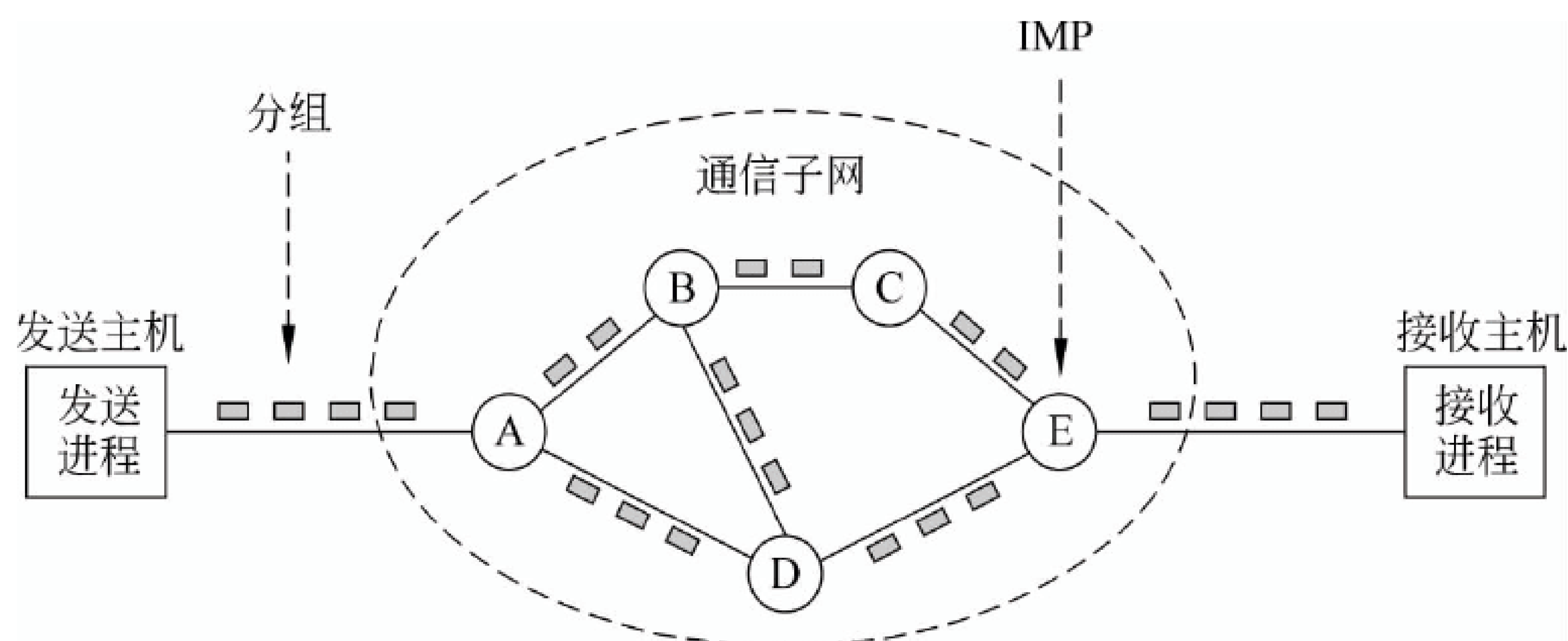


图 1-10 分组交换网结构与原理示意图

ARPANET 的设计目标如下。

(1) 传输计算机的数字数据信号。

(2) 网络可以连接不同型号的计算机。

(3) 网络中所有的节点都是同等重要的。

(4) 网络必须有冗余的路由。

(5) 网络结构简单,必须保证数据的正确传输。

根据 ARPA 提出的设计要求,ARPANET 在总体方案中采取了分组交换(Packet Switches)的思想。早期的 ARPANET 实际上是一个广域计算机网络。ARPANET 由通信子网与资源子网两个部分组成。通信子网的报文存储转发节点由一些小型计算机组成,这些小型计算机被称作“接口报文处理器(IMP)”。

ARPA 是以招标的方式来建立通信子网的,当时一共有 12 家公司、研究部门与大学参与了竞标。ARPA 将项目建设内容分为三项,经评估后有三家赢得了合同,他们的分工是:BBN(Bolt Beranek & Newman)公司与 Frank Heart 承担分组交换关键设备 IMP 的研发;网络分析公司(Network Analysis Corporation)与 Roberts、Howard Frank 负责网络拓扑结构的设计、优化和网络经济性的研究任务;UCLA 与 Leonard Kleinrock 负责网络测试中心的建设任务。

BBN 公司在通信子网的组建中,选择了 Honeywell 公司的 DDP 316 小型计算机(内存为 16b、12KB)作为 IMP,这些小型计算机都是经过特殊改进的。由于考虑到计算机系统的可靠性,IMP 没有采用外接磁盘系统。出于经济上的原因,当时通信线路租用的是电话公司的 56kbps 线路。图 1-11 是作为第一台 IMP 的 DDP 316 小型计算机的照片。

最初实验网络的每个节点都由一台 IMP 和一台主机组成,它们位于同一个房间中,并且通过一条很短、速率为 56kbps 的电缆连接起来。主机给 IMP 发送的报文最长为 8063b,IMP 把报文分成长度为 1008b 的分组,再独立地将这些分组向下一个节点转发。下一个 IMP 完整地接收到一个数据分组之后存储起来,检查传输过程中没有出错,再向它的下一



图 1-11　作为第一台 IMP 的 DDP 316 小型计算机

个 IMP 节点转发,直至到达目的 IMP。目的 IMP 将正确接收的属于同一个文件的分组重新组装成报文后,再递交给与它直接连接的目的主机。这样,一个分组“存储转发”的过程就结束了。

在实验过程中,为了保证网络通信系统的高度可靠性,要求每个 IMP 都至少连接到两个其他的 IMP。如果有某些 IMP 设备或通信线路被毁坏,仍然可以通过网络中其他的路径,自动完成分组的转发。IMP 就是目前广泛使用的路由器的雏形。

当时在 BBN 公司任职的 Robert Kahn(罗伯特・卡恩)在 IMP 与 TCP/IP 的研究中发挥了重要的作用。鉴于 Robert Kahn 在 ARPANET 的 IMP,以及 TCP/IP 与 Internet 研究中的突出贡献,Robert Kahn 与 Vinton G. Cerf 一起被誉为“Internet 之父”,并获得 2004 年的图灵奖(如图 1-12 所示)。

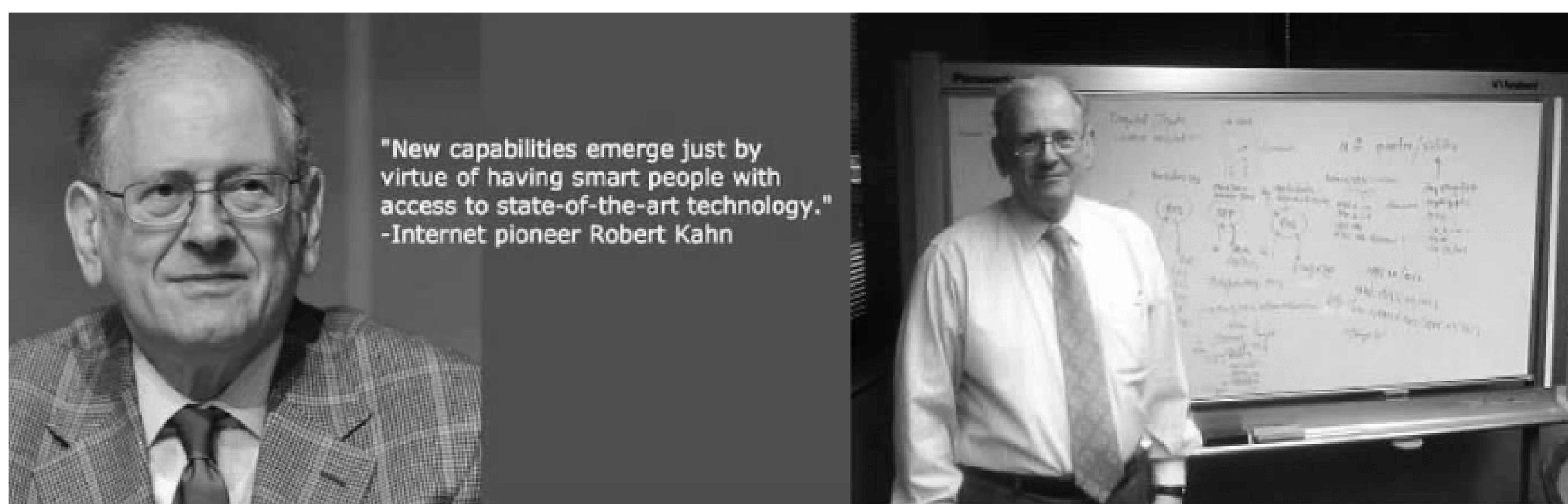


图 1-12　2004 年图灵奖获得者——Robert Kahn

2. ARPANET 协议与网络软件结构

ARPANET 通信协议采用的是网络控制协议(Network Control Protocol,NCP)。1969 年 4 月 7 日,Steve Crocker 发表了 RFC1,题为 “Host Software”的文档。NCP 协议就是在它的基础上发展起来的。

ARPANET 软件包括两个部分:执行子网内部通信协议的软件和执行主机—主机通信协议的软件。开发网络软件首先要制定网络协议。子网内部协议主要包括 HOST-IMP 与 IMP-IMP 的通信协议,还需要专门设计用来提高从源 IMP 到目的 IMP 传输可靠性的通信协议。实现子网内部通信协议的软件包括:HOST-IMP 通信软件与 IMP-IMP 通信软件。实现主机端通信协议的软件包括:HOST-IMP 的通信软件、HOST-HOST 的应用软件。第一个 ARPANET 实验系统使用的主机到主机通信的通信协议是网络控制协议(NCP)。图 1-13 给出了 ARPANET 的协议结构示意图。

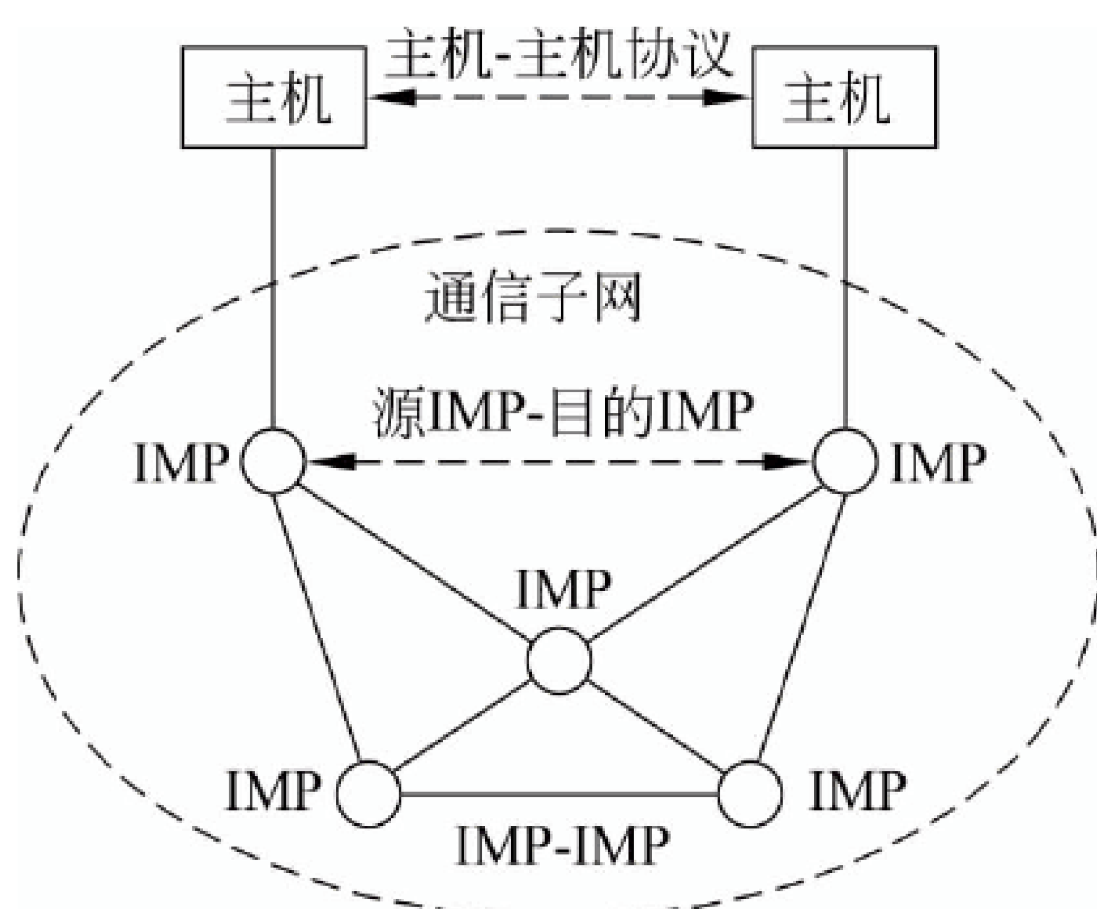


图 1-13　ARPANET 协议结构

3. ARPANET 软件研发与网络组建的过程

在完成网络结构与硬件设计后，一个重要的问题是需要开发软件。1969 年夏季，Larry Roberts 在犹他州的 Snowbird 召集网络研究人员会议，参加会议的大多数是研究生。研究生们希望像完成其他编程任务一样，有网络专家向他们解释网络的设计方案与需要编写的软件，然后分配给每人一个具体的软件编程任务。当他们发现那里没有网络专家，也没有完整的设计方案时很吃惊。他们必须自己想办法找到自己该做的事情。

1969 年 12 月，包含 4 个节点的实验网络开始运行，这 4 个节点是 UCLA（加州大学洛杉矶分校）、UCSB（加州大学圣塔芭芭拉分校）、SRI（斯坦福研究院）和 University Utah（犹他大学）4 所大学。选择这 4 所大学是由于它们都与 ARPA 签订了合同，而且都有不同类型的主机。

节点 1（UCLA）是在 1969 年 8 月 30 日～9 月 2 日期间接入；节点 2（SRI）是在 1969 年 10 月 1 日接入；节点 3（UCSB）是在 1969 年 11 月 1 日接入；节点 4（University Utah）是在 1969 年 12 月接入。

第一台 IMP 安装在 UCLA，其他的三台分别安装在 UCSB、SRI 与 University Utah。据当时负责安装第一台 IMP 的 UCLA 计算机系教授 Leonard Kleinrock（伦纳德·克兰罗克）回忆，1969 年 9 月 2 日第一台 IMP 安装调试成功。1969 年 10 月 1 日第二台 IMP 在 SRI 安装，第一个数据分组成功地从 UCLA 发送到 SRI。为了验证数据传输的情况，参加实验的双方使用了语音通话设备来相互联系。1969 年 10 月 29 日 22:30，Leonard Kleinrock 让研究人员从 UCLA 远程登录到 SRI 主机时，在输入由 5 个英文字母组成的登录命令“LOGIN”中的“L、O”两个字母后，网络系统出现了故障，第一次远程登录失败。但是，这是一个非常重要的时刻，它标志着计算机网络时代已经开始到来。很多人认为应该将这一天作为 ARPANET，以及由它发展起来的 Internet 的诞生日。图 1-14 给出了 ARPANET 最初的 4 个节点的结构示意图、实验记录与 Leonard Kleinrock 教授的照片。

1969 年，Leonard Kleinrock 在向新闻界发表谈话时说：“一旦 ARPANET 建立并运行起来，我们从家中和办公室访问计算机系统，就像我们获得电力或电话服务那样容易。”现在读到这段话时，读者会发现 1969 年 Leonard Kleinrock 的预见与现在研究的“普适计算”“云计算”的概念是如此的吻合。

从 1969 年到 1971 年，经过近两年对网络应用层协议的研究与开发，研究人员首先推出了 TELNET 应用。1971 年 2 月公布了第一个关于 TELNET 协议的文档 RFC97。1972 年，



图 1-14　ARPANET 最早的 4 个节点的结构草图

ARPANET 节点数增加到 15 个。1972 年 10 月，Robert Kahn 在华盛顿 DC 召开的第一届国际计算机与通信会议(ICCC)上做了有关 ARPANET 首次的公开演示。当时参加演示的 40 台计算机分布在美国各地，演示的项目包括网上聊天、网上弈棋、网上测验、网上空管模拟等，其中，网上聊天演示引起了极大的轰动，吸引了世界各国计算机与通信学科的科学家加入到计算机网络研究的队伍之中。图 1-15 给出了 1972 年 4 月的 ARPANET 拓扑结构示意图。

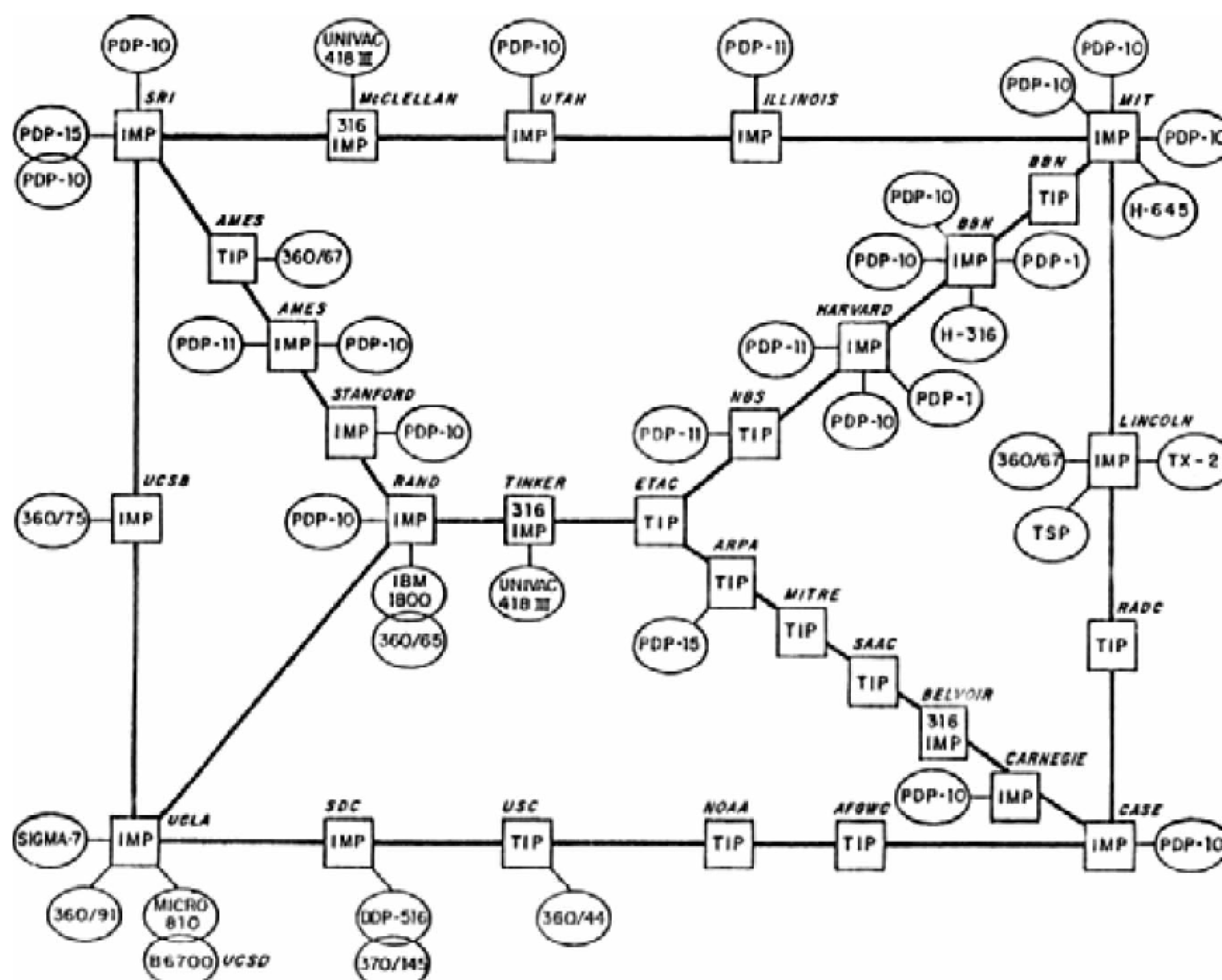


图 1-15　1972 年 ARPANET 拓扑结构示意图

随着英国伦敦的大学节点与挪威的皇家雷达研究所节点接入 ARPANET,使得 ARPANET 的节点数增加到 23 个,同时也标志了 ARPANET 已经国际化。

1972 年,BBN 的 Ray Tomlinson 编写第一个用于网络的电子邮件 E-mail 应用程序,当时接入 ARPANET 的节点数大约在 40 个。1973 年,E-mail 的通信量已占到 ARPANET 总通信量的 3/4。随着更多的 IMP 被交付使用,ARPANET 规模快速地增长起来,很快就扩展到了整个美国。

4. ARPANET 对推动计算机网络技术发展的贡献

ARPANET 是一个典型的广域网系统,它的研究成果标志着广域网技术的成熟,并且进入应用阶段。同时,它在推动计算机网络理论与技术的发展上有着深远的意义。

分组交换的概念最初是在 1964 年提出。1969 年 12 月,美国第一个使用分组交换技术的网络 ARPANET 投入运行。人们认为,分组交换技术的出现标志着现代电信时代开始。ARPANET 是计算机网络技术发展的一个重要的里程碑,它对计算机网络理论与技术发展起到了重大的奠基作用。它的贡献主要表现在以下几个方面。

(1) 完成了对计算机网络定义与分类方法的研究。

(2) 提出了计算机网络体系结构与参考模型的概念。

(3) 研究并实现了分组交换方法。

(4) 完善了层次型网络体系结构的模型与协议体系。

(5) 开始了 TCP/IP 模型、协议与网络互联技术的研究与应用。

到 1975 年,ARPANET 中已经连入一百多台主机,并且结束网络实验阶段,移交给美国国防部国防通信局正式运行。

1983 年 1 月,ARPANET 向 TCP/IP 的转换全部结束。根据美国法律,所有政府出资的项目应体现纳税人的权利,都必须由纳税人分享。因此,由国防部出资研究的 ARPANET 必须允许与国防无关的其他大学与研究部门科研人员使用。出于对军事机密的需要,美国国防部国防通信局将 ARPANET 分成两个独立的部分:一部分仍叫作 ARPANET,继续用于一般的科学研究工作,成为后来的 Internet;另一部分稍大一些的 MILNET,用于军方的非机密通信(如图 1-16 所示)。

20 世纪 80 年代中期,随着 ARPANET 的规模不断增大,ARPANET 成为 Internet 的主干网。1990 年,ARPANET 已被新的网络所替代。虽然 ARPANET 目前已经退役,但是人们将会永远记住它,这是因为它对网络技术的发展产生了重要的影响。到目前为止,MILNET 仍然在运行着。

20 世纪 70 年代到 20 世纪 80 年代,网络技术发展十分迅速,并且出现大量的计算机网络,仅美国国防部就资助建立了多个计算机网络。同时,还出现了一些研究实验性网络、公共服务网络和校园网。在这个阶段中,公共数据网(Public Data Network,PDN)发展迅速。所谓"公共数据网"是指由专门的网络运营商(Network Carrier)或网络服务提供商(Network Service Provider,NSP)运营与管理的网络。传统的电信网络主要是提供电话语音通信服务。在计算机网络应用的初期,网络运营商大多是电话公司,它们在提供电话服务的同时,也提供专线服务。而网络服务提供商有不同的类型,它们可能没有自己的网络基础设施,而是租用电话公司的线路,提供数据通信服务。在互联网高速发展阶段,网络运营商与网络服务提供商纷纷转向提供本地互联网用户接入服务,成为互联网服务提供商 ISP。

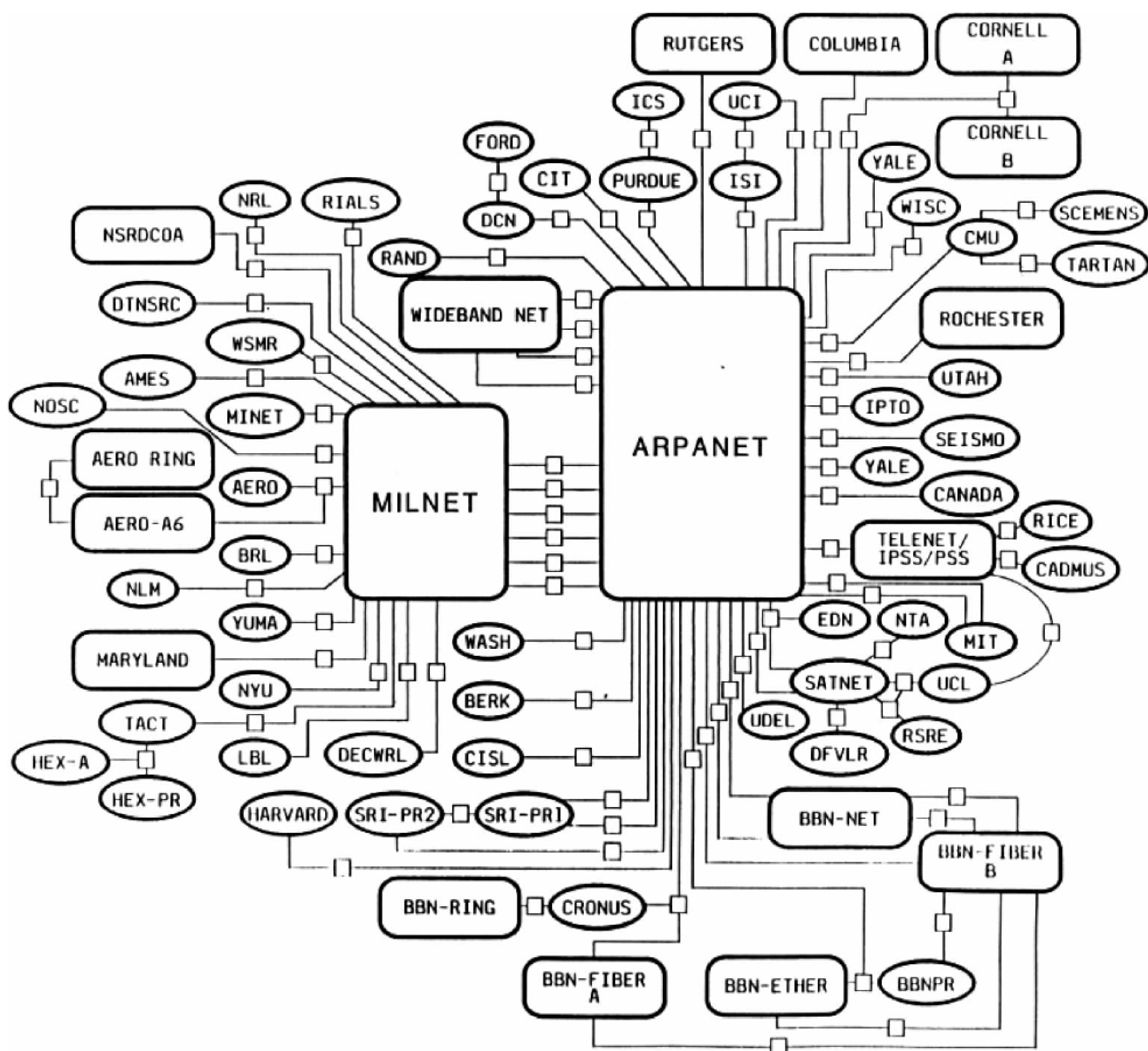


图 1-16　ARPANET 分成两个独立的部分

问题 1-3：TCP/IP 经历了怎样的发展与演变过程？

现在术语“TCP/IP”与互联网一起被广大网民所接受，这也说明了 TCP/IP 的重要性。讲授互联网必须懂得 TCP/IP，而深入理解 TCP/IP 发展与演变的过程，就需要注意以下几个基本的问题。

1. TCP/IP 产生的背景

1972 年，ARPANET 的核心研究人员 Vinton G. Cerf(温顿·瑟夫)与 Robert Kahn(罗伯特·卡恩)开展了“网络互联项目”的研究。他们希望将不同类型的网络互联起来，使不同类型网络中的主机之间可以互联互通。网络互联需要克服异构网络在分组长度、分组结构、报头结构与传输速率等方面的差异。Vinton G. Cerf 与 Robert Kahn 提出用一种称为网关(Gateway)的设备实行网络的互联。实际上，当时提出的网关从功能上来说仍是现在使用的路由器。

1973 年，ARPA 资助了分组无线网络(Packet Radio Netork，PRNET)与分组卫星网络(Packet Satellite Network，SANET)的研究。Robert Kahn 最初考虑在继续使用 NCP 的基础上，开发只适合无线分组网络的通信协议。在实际研究工作中，Robert Kahn 发现：NCP 寻址功能只能够做到将分组传送到最后一个 IMP，并且 NCP 不提供端-端的差错控制能力。NCP 是依靠高层的应用程序去保证端-端分组传输的可靠性，一旦一个分组丢失就有可能

造成死机。针对 NCP 的不足，Robert Kahn 认为：互连不同类型计算机的关键是需要建立一个开放的网络结构和可靠的端到端协议。Robert Kahn 提出了互联网络结构与网络互联协议设计原则，可以总结如下。

（1）网络与网络互连时不需要改变内部的结构。

（2）互连的网络通信采取“尽力而为”的原则。

（3）互联网络的路由器与网关不保留转发的分组。

（4）不存在全网集中控制的中心节点。

Robert Kahn 提出了网络互联协议设计原则充分体现出：“分布、平等、协作、服务”的特点。正是在这种正确的网络协议设计原则指导下，才使得 ARPANET 与 Internet 能够顺利和快速地发展。

1974 年 5 月，Vinton G. Cerf 与 Robert Kahn 在 *IEEE Transactions on Communications Technology* 上发表了具有里程碑意义的论文。这篇论文阐述了实现分组端-端交付的传输控制协议 TCP，论文内容涉及分组封装、报头结构与网关协议等关键问题的解决方案。

1977 年 10 月，ARPANET 完成了与分组无线网络、分组卫星网络互联的实验。通过实验，他们决定将初期的 TCP 分成两个协议，即现在使用的传输控制协议 TCP 与互联网络协议 IP。互联网络协议 IP 负责处理分组路由，而传输控制协议 TCP 则负责报文的分段、重装、差错检测与进程通信管理。这项研究工作奠定了目前广泛使用的 TCP/IP 的基础。

在了解 TCP/IP 体系与 Internet 发展时，需要介绍对此做出重要贡献的 Vinton G. Cerf 教授。Vinton G. Cerf 在斯坦福大学获数学学士学位，在 UCLA 获计算机博士学位。在 UCLA 读研究生阶段，他参与早期的 ARPANET 测试与软件编程工作。在 UCLA 安装第一台 IMP 时，他就在这个实验室工作。1976—1982 年，Vinton G. Cerf 在 ARPA 任职期间，负责 ARPANET 协议设计的研究工作。这段时间的主要工作是研究 TCP/IP 和设计 IP 地址体系。1982—1986 年，Vinton G. Cerf 作为 MCI 数字信息服务公司副总裁，领导了世界上第一个 Internet 商用电子邮件服务系统的研究与运行。1986 年后，Vinton G. Cerf 作为斯坦福大学教授，一直领导 TCP/IP 软件与路由器的研发工作。图 1-17 是 2004 年图灵奖获得者 Vinton G. Cerf 的照片。



图 1-17　2004 年图灵奖获得者 Vinton G. Cerf

2005 年 2 月 16 日，美国计算机协会 ACM 宣布 Vinton G. Cerf 和 Robert Kahn 获得

2004年图灵奖,以表彰他们在设计和实施Internet通信协议方面的重大成就(如图1-18所示)。ACM主席认为,他们在Internet体系结构和协议的研究中取得的成果是奠定网络技术的基石。“他们的工作使我们今天依赖的许多快速方便的应用成为可能,其中包括电子邮件、Web、实时通信、P2P传输、协同工作与会议工具等。这些发展帮助信息技术成为整个行业的关键部分。”这项图灵奖也向世人说明了计算机网络技术在信息技术中的“基石”作用,确定了Internet技术的学术地位。

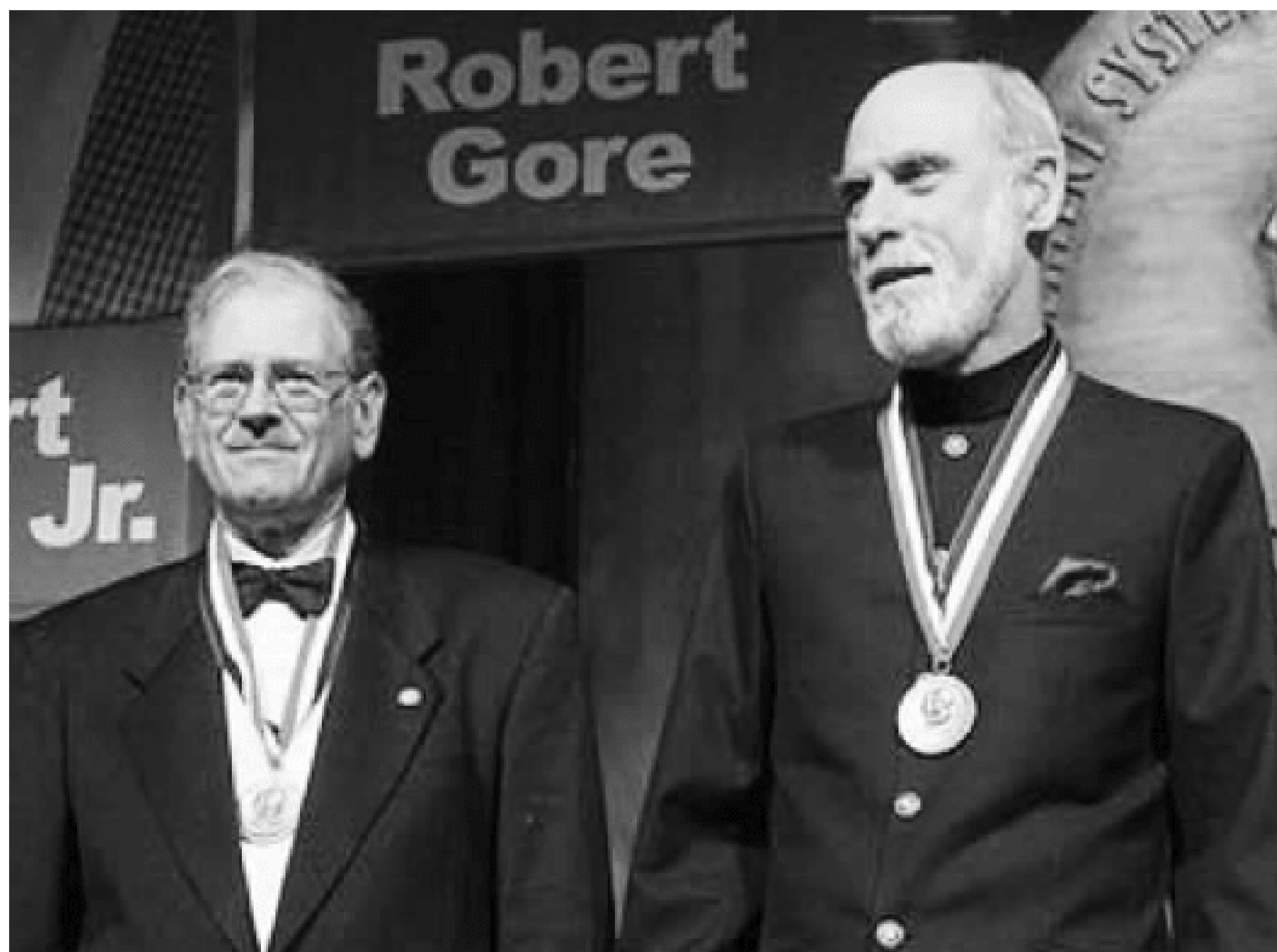


图1-18　Vinton G. Cerf与Robert Kahn

2. TCP/IP的应用

1978年,TCP/IP可以真正运行的版本宣告研发成功。1981年,美国ARPA决定选择TCP/IP作为军方网络的协议标准;1982年,ARPA决定将ARPANET上所有的系统全部从NCP转换为TCP/IP。为了保证协议转换工作顺利进行,ARPA决定在数据链路层为传输NCP与IP分组分配不同的信道,从而使中止NCP时不至于引起很大的混乱。但是,1982年ARPANET关闭了NCP一天,结果引起了一片混乱,因为很多计算机都没有启动TCP/IP。1982年秋天,NCP又被关闭了两天。1983年1月1日,NCP被彻底关闭,ARPANET全部运行TCP/IP。

随着越来越多的网络接入ARPANET,网络互联也变得越来越重要。为了鼓励采用TCP/IP,ARPA、BBN公司和加州大学Berkeley分校签订了合同,将新的TCP/IP集成到Berkeley UNIX操作系统中。Berkeley的研究人员开发了一个方便的、专门用于连接网络的编程接口,并编写了很多应用程序、开发工具与管理程序,这些工作使得网络互联变得更容易。1983年,BSD UNIX 4.2操作系统正式推出,很多大学采用了BSD UNIX,这项工作也促成TCP/IP的普及。

BSD UNIX在网络方面成功的原因主要是提供标准的TCP/IP应用程序,以及一组网络服务工具程序。这些工具与UNIX命令的调用方式相似,因此受到广大UNIX用户的欢迎。BSD UNIX提供了可以访问操作系统的编程接口的应用程序,使程序员可以方便地访问TCP/IP。同时,SUN公司将TCP/IP引入了商业领域。

TCP/IP的成功促进了Internet的发展,Internet发展又进一步扩大了TCP/IP的影响。IBM、DEC等大公司纷纷宣布支持TCP/IP,各种网络操作系统与大型数据库产品开始支持TCP/IP。随着Internet的广泛应用和高速发展,TCP/IP体系结构已成为业内公认的

标准。TCP/IP 的发展演变过程如图 1-19 所示。

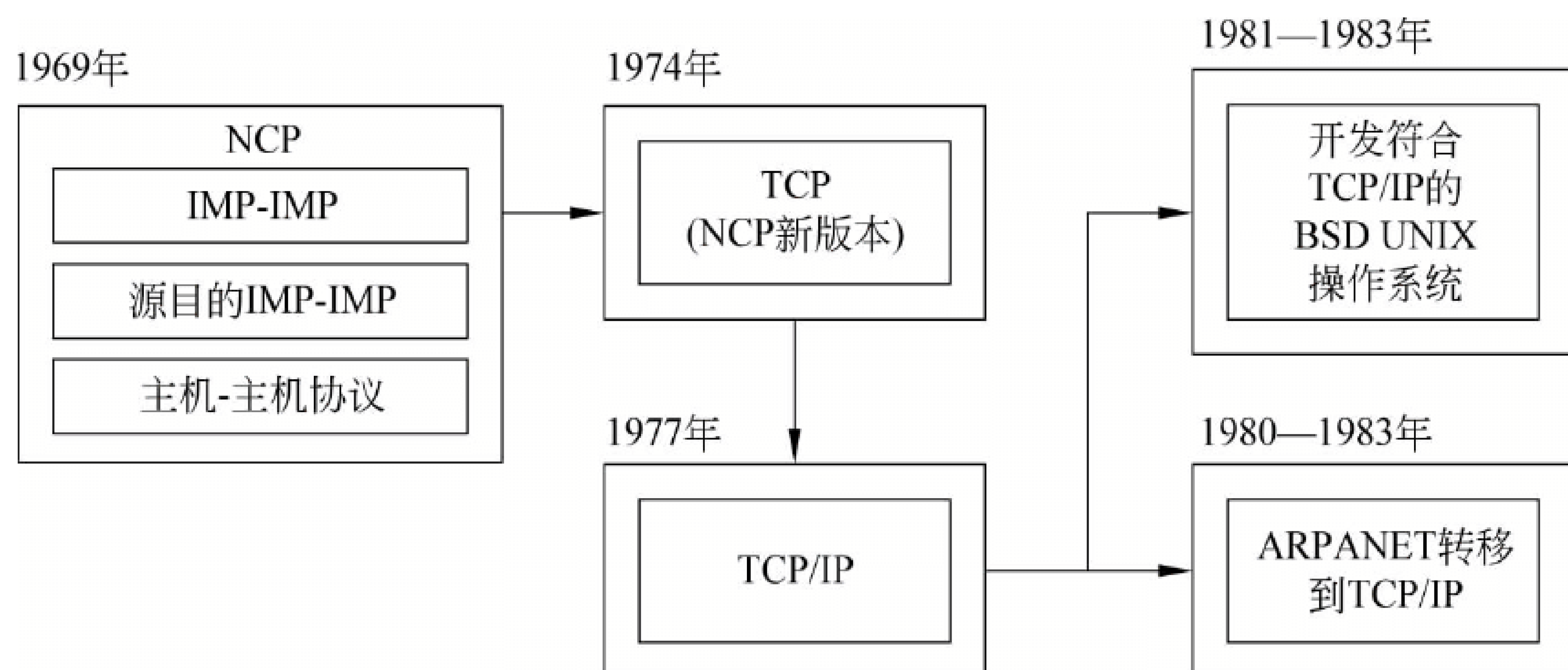


图 1-19　TCP/IP 的发展演变过程示意图

问题 1-4：如何认识互联网形成与发展的过程？

了解互联网形成与发展的过程对讲授网络技术有很大的帮助。了解互联网形成与发展过程需要注意以下几点。

1. NSFNET 对 Internet 发展的影响

20 世纪 70 年代后期，随着 ARPANET 在美国高校中的应用不断地扩大，美国国家科学基金会(NSF)认识到 ARPANET 对研究工作的重要影响。但是，不是所有大学都有这样的机会，连入 ARPANET 的大学必须与美国国防部有合作研究项目。20 世纪 80 年代，个人计算机与工作站开始普及，各种局域网(如 Ethernet、Token Bus、Token Ring、NetWare 等)技术日趋成熟，促进了 Internet 的发展。同时，美国 NASA 为空间科学家建设的 SPAN 网络、美国能源部为能源研究建设的 HEPNET，以及连接美国各个大学 IBM 中心大型计算机的 BITNET，只能通过网关与 Internet 交换电子邮件。

1981 年，为了使更多的大学能够共享 ARPANET 的资源，NSF 计划建设一个虚拟网络，即计算机科学网(Computer Science Network，CSNET)。CSNET 的中心是一台 BBN 计算机，不能直接连入 ARPANET 的大学可以通过电话拨号与 BBN 计算机连接，通过这台 BBN 计算机作为网关，间接连入 ARPANET。1981 年，CSNET 接入到 ARPANET，它连接了美国所有大学的计算机系。

接入 ARPANET 的主机数剧增促进了域名技术的发展。随着 TCP/IP 的标准化，ARPANET 的规模一直在不断扩大，不仅美国国内有很多网络与 ARPANET 相连，世界上很多国家也通过远程通信线路，采用 TCP/IP 将本地的计算机与网络连入 ARPANET。针对 TCP/IP 互联的主机数量急剧增加的情况，网络中计算机的管理和系统的运行成为迫切需要解决的问题。在这种背景之下人们提出了域名系统(Domain Name System，DNS)。DNS 将多个主机划分成不同的域，通过域名来管理和组织互联网中的主机。DNS 使用分布式数据库存储与主机命名相关的信息。域名系统使得在物理结构“无序”的网络，变成从逻辑结构上“有序”的、可管理的系统。最初记录主机名与 IP 地址对应关系的是一个静态的文本文件 HOSTS。到 1982 年，人们发现用简单的文本文件去记录所有联网的主机名与 IP 地址已越来越困难。Paul Mockapetris 着手设计一个分布式数据库系统，即 DNS。1984 年，第一个 DNS 程序 JEEVES 开始使用。1988 年，BSD UNIX 4.3 推出了它的 DNS

程序 BIND。

1984 年,NSF 决定组建 NSFNET。NSFNET 的主干网连接美国 6 个超级计算机中心,它们分布在 San Diego、Boulder、Champaign、Pittsburgh、Ithaca 和 Princeton。NSFNET 的通信子网使用的硬件与 ARPANET 基本相同,采用的是 56kbps 的通信线路。但是,NSFNET 的软件技术与 ARPANET 不同,它从开始就使用了 TCP/IP,成为第一个使用 TCP/IP 的广域网。

1990 年,ARPANET 停止了运行,NSFNET 代替 ARPANET 成为 Internet 的主干网。NSFNET 采用的是一种层次型结构,分为主干网、地区网与校园网。各大学的主机连入校园网,校园网连入地区网,地区网连入主干网,主干网再通过高速通信线路与 ARPANET 连接。包括主干网与地区网在内的整个网络系统称为 NSFNET。连入校园网的主机用户可以通过 NSFNET,访问任何一个超级计算机中心的资源,访问与网络连接的数千所大学、研究实验室、图书馆与博物馆,用户之间相互交换信息、发送和接收电子邮件。

由于 NSFNET 鼓励和资助很多大学、政府机构、研究部门接入网络,因此 1986—1991 年 NSFNET 的子网个数从 100 个迅速增长到 3000 个,几乎是以每年 100%的速度增长。

在 NSFNET 迅速扩张的同时,出现了网络负荷过重的情况,NSF 决定立即开始研究下一步发展策略。随着网络规模的继续扩大和应用的扩展,NSF 认识到政府已经不能继续从财政上支持这个网络。虽然有不少商业机构打算参与进来,但是 NSF 并不允许这个网络用于商业用途。在这种情况下,NSF 鼓励 MERIT、MCI 与 IBM 等三家公司组建一个非营利性的公司运营 NSFNET。MERIT、MCI 与 IBM 三家公司合作创建了 ANS 公司。

1990 年,ANS 公司接管了 NSFNET,并在全美范围内组建了 T3 级的主干网,网络传输速率为 44.746Mbps。到 1991 年年底,NSFNET 的全部主干网节点都与 T3 主干网连通。图 1-20 给出了 1991 年 9 月 NSFNET 主干网与节点分布示意图。

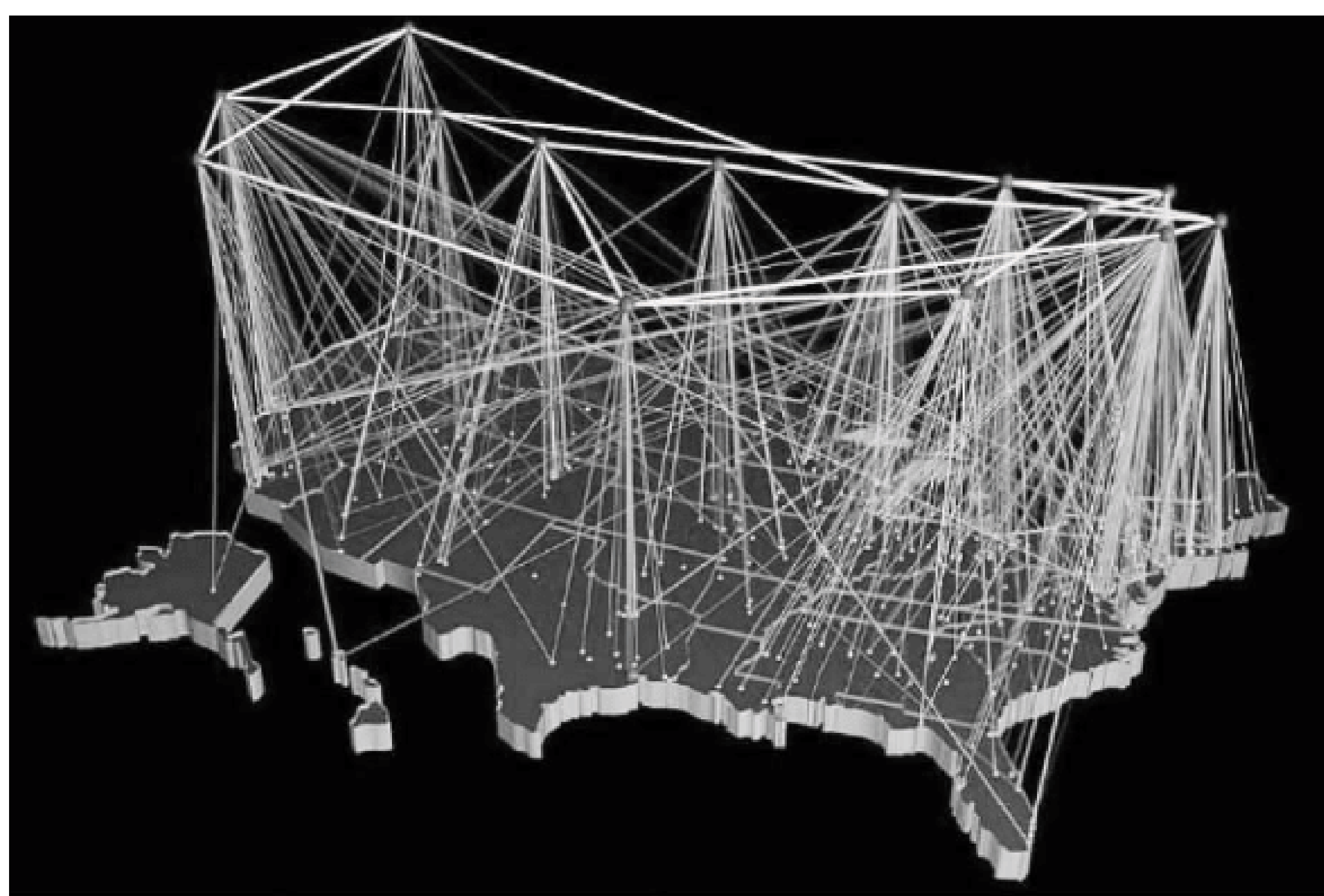


图 1-20　1991 年 9 月 NSFNET 主干网与节点分布示意图

在美国发展 NSFNET 的同时,其他国家与地区也在建设与 NSFNET 兼容的网络,例如,欧洲为研究机构建立的 EBONE、Europa NET 等。当时,这两个网络都采用了 2Mbps 的通信线路与欧洲很多城市连接。欧洲每个国家都有一个或多个国家网,它们都与 NSFNET 的地区网兼容。这些网络为 Internet 的发展奠定了基础。

1991 年，由于 NSF 只支付 NSFNET 主干网 10% 的通信费，因此 NSF 开始放宽对 NSFNET 使用的限制，允许商业信息通过 NSFNET 主干网传输。随着 Internet 私营化的出现，1995 年 Internet 完全实现了私有化，NFC 不再向 Internet 的主干网提供资金。NSFNET 主干网成为 Internet 的主干网。NSF 在 1995 年 4 月 30 日正式将 NSFNET 退役。美国大部分主干网业务开始由商业互联网交换中心（CIX）互联的互联网服务提供商 IPS 提供服务。互联网完全实现了私有化。

美国 NFC 从 1986 年建立 NSFNET 到 1995 年 Internet 私有化，这段时间一共资助了两亿美元。此时，美国接入 Internet 的网站约为 2.9 万个，全世界接入的网站超过了 5 万个。图 1-21 给出了从 ARPANET 到 Internet 的发展过程。

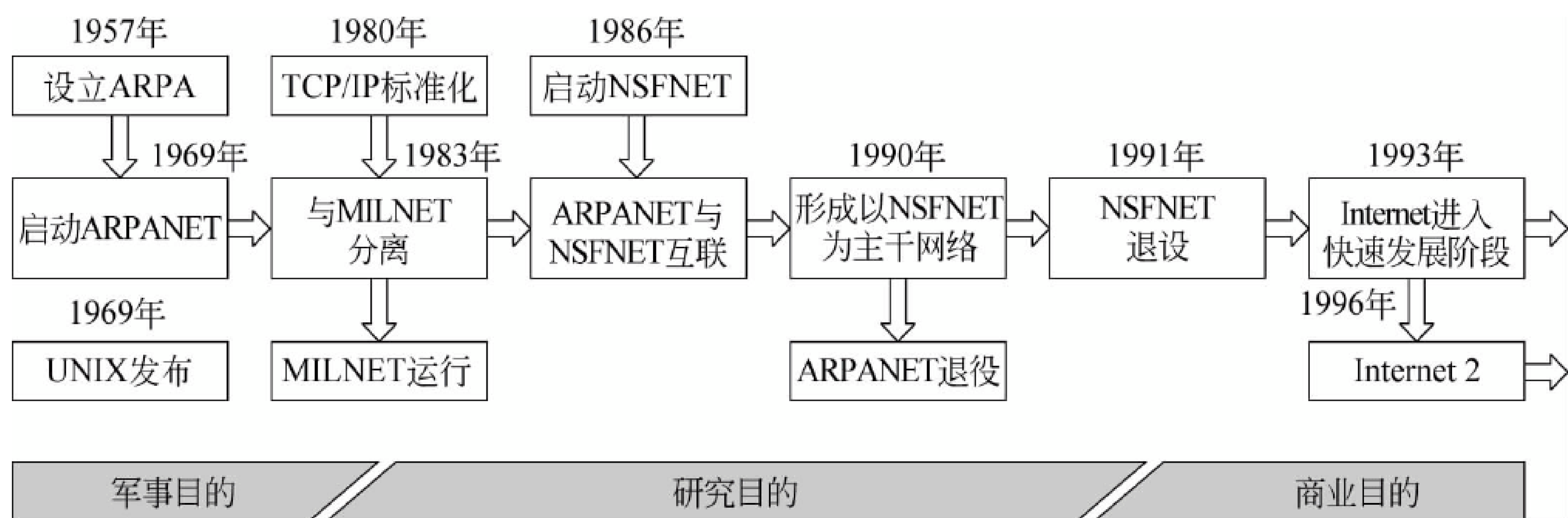


图 1-21　从 ARPANET 到 Internet 的发展过程

1995 年 4 月，NSF 和 MCI 开始合作建设高速主干网（very High Speed Backbone Network Service，vBNS）。vBNS 用来代替原有的 NSFNET 主干网。尽管 vBNS 是 NSF 和 MCI 合作的成果，但是它的应用仍被限制在 NSF 所批准的教育和研究机构。后来，vBNS 主干网的传输速率从 622Mbps（OC12）提高到 4.8Gbps（OC48）。

2. Internet 的形成

1983 年 1 月，TCP/IP 正式成为 ARPANET 的网络协议标准。此后，大量的网络、主机与用户连入 ARPANET，使得 ARPANET 得到迅速发展。随着很多地区性网络连入 ARPANET，这个网络逐步扩展到其他国家与地区。很多现存的网络都已经连入 Internet，包括空间物理网（SPAN）、高能物理网（HEPNET）、IBM 的大型计算机网络（BITNET）与西欧的欧洲学术网等。20 世纪 80 年代中期，人们开始认识到这种大型互联网络的作用。20 世纪 90 年代是 Internet 历史上发展的黄金时期，其用户数量以平均每年翻一番的速度增长。

Internet 的最初用户只限于科学研究和学术领域。20 世纪 90 年代初期，Internet 上的商业活动开始缓慢发展。1991 年，美国成立了商业网络交换协会，允许在 Internet 上开展商务活动，各个公司逐渐意识到 Internet 在宣传产品、开展商贸活动上的价值，Internet 上的商业应用开始迅速发展，其用户数量已超出学术研究用户一倍以上。传统的 Internet 应用主要有 FTP、E-mail、TELNET 与 Gopher 等。商业应用的推动使 Internet 的发展更加迅猛，规模不断扩大、用户不断增加、应用不断拓展、技术不断更新，使 Internet 几乎深入到社会生活的每个角落，成为一种全新的工作方式、学习方式和生活方式。

目前，ANS 公司建设的 ANSnet 是 Internet 主干网，其他国家或地区的主干网都通过

ANSnet接入Internet。家庭用户通过电话线连接到Internet服务提供商(Internet Service Provider,ISP)。办公室的计算机通过局域网连入校园网或企业网。局域网分布在各个建筑物内,连接各个系所与研究室的计算机。校园网、企业网通过专用通信线路与地区网络连接。校园网中的各种主机都是用户可以访问的重要资源。这些系统都通过校园网连入Internet,供本校或其他大学的Internet用户访问。

从用户的角度来看,Internet是一个全球范围的信息资源网,接入Internet的主机可以是信息服务提供者的服务器,也可以是信息服务使用者的用户计算机。Internet代表着全球范围内无限增长的信息资源,是人类拥有的最大的知识宝库之一。随着Internet规模的扩大,网络与主机数量的增多,它能提供的信息资源与服务将会更加丰富。

1995年10月24日,在广泛征求意见的基础上,美国联邦网络委员会通过一项提案,为Internet做出如下的定义:“Internet是一个全球性的信息系统,系统中的每一台主机都有一个全球唯一的主机地址,IP协议定义了地址的格式。系统中主机与主机之间的通信遵守TCP/IP,或者是其他兼容的协议来交换信息。在以上描述的信息基础设施之上,利用公网或专网的形式,向社会大众提供资源与服务。”

问题1-5:如何认识Web技术对互联网应用发展的影响?

如果说开放网络服务是促进Internet快速发展的第一次飞跃的推动力,那么Web技术的出现就是Internet第二次快速发展的推动力。

1. Tim Berners-Lee与Web发明的过程

20世纪80年代后期超文本技术已经出现,当时就有专门讨论超文本(Hypertext)的国际学术会议,每次会议都会有上百篇的论文,但没有人想到可以将超文本概念应用到Internet的信息共享之中。Web技术于1989年诞生于欧洲原子能研究中心CERN。它是从一种用于分发高能物理数据的方法,发展成为Internet的重要的应用。

欧洲原子能研究中心CERN有几台加速器,来自欧洲各个国家的科学家利用这些加速器来进行粒子物理研究。这些实验一般需要几年的时间进行准备。参加研究的80个国家的6500名科学家与工程师,需要不断地通过欧洲原子能研究中心CERN交换文件。其中也包括中国的科学家。当时交换文件的过程很不方便,大家对这个方法提出了很多批评。如何方便地为参加研究的科学家提供交换文件的方法成为一个重要的研究课题,Web技术正是源于这样一个实际的应用需求。

1989年3月,欧洲原子能研究中心CERN的Tim Berners-Lee(蒂姆·伯纳斯-李)提出用超文本技术链接HTML文档的建议。他试图让这些分散在各个国家的研究人员,通过交换报告、计划、图纸、照片等各种文档来方便地开展合作研究。他工作的重点是解决HTML、HTTP、Web服务器与浏览器设计等4项关键技术。

1976年,Tim Berners-Lee从牛津大学毕业,获物理学学士学位。由于他的父母都从事于计算机研究工作,因此他对计算机技术产生了浓厚的兴趣。1989年,Tim Berners-Lee进入日内瓦的欧洲粒子物理实验室CERN工作。在这里Tim Berners-Lee接受了一项极富挑战性的工作,CERN委托他开发一个软件,目的是让欧洲各国的核物理学家能通过计算机网络,方便地共享最新的实验数据与图像资源,开展合作研究。这个网络应用软件的开发激发了Tim Berners-Lee的研究热情,他将进一步研究的目标瞄向了建立一个全球范围的信息网络共享的研究上,并于1989年3月向CERN递交了一份立项建议书,建议采用超文本

技术把欧洲原子能研究中心 CERN 内部的各个实验室连接起来，并将这项研究成果推广到全世界。这个激动人心的建议在 CERN 内部引起轩然大波，但这里终究是核物理实验室，并非计算机网络研究中心。尽管有人支持这项研究工作，但是最终申请还是没有被批准。Tim Berners-Lee 并没有灰心，他花了两个月重新修改了建议书，加入了对超文本系统开发步骤与应用前景的阐述，再一次呈报后终于得到了批准。Tim Berners-Lee 得到了一笔研究经费，购买了一台 NEXT 计算机，并率领助手开发实验系统。1989 年 12 月，Tim Berners-Lee 为这个系统定名为“World Wide Web”，即“WWW”或“Web”。

1990 年 9 月，第一个基于文本链接的原型系统（ENQUIRE）投入运行。他建立的第一个网站 http://info. cern. ch/于 1991 年 8 月 6 日正式运行。虽然当时这个 Web 服务器功能十分简单，看起来像是 CERN 的电话号码簿，它只是允许用户进入主机来查询每个研究人员的电话号码，但是它向用户提供了一个充分体现超文本概念的浏览器。1991 年 5 月，WWW 在 Internet 上首次露面，立即引起轰动。尽管 Tim Berners-Lee 在 2009 年 10 月 15 日的一次学术会议上回忆当时设计 WWW 网址时坦承：当初在网址前面加上双斜线“//”实在有欠考虑。当时没想那么多，也没想到今天 WWW 应用会如此广泛。当时只想到要用个符号“告诉”计算机接下来输入的是网站地址，所以想到了用斜线这个不常用的标点。

在 1991 年 12 月 Texas 的 San Antonio 举行的 Hypertext 91 会议上，Tim Berners-Lee 对 Web 技术进行了一次公开演示。这次演示引起了很多研究人员的注意。图 1-22 是 Web 技术的发明者 Tim Berners-Lee、NEXT 计算机与 WWW 界面、报告文档，以及他在 1991 年学术会议上向代表们演示 WWW 应用的照片。WWW 技术推动了 Internet 应用发展的第二次高潮。

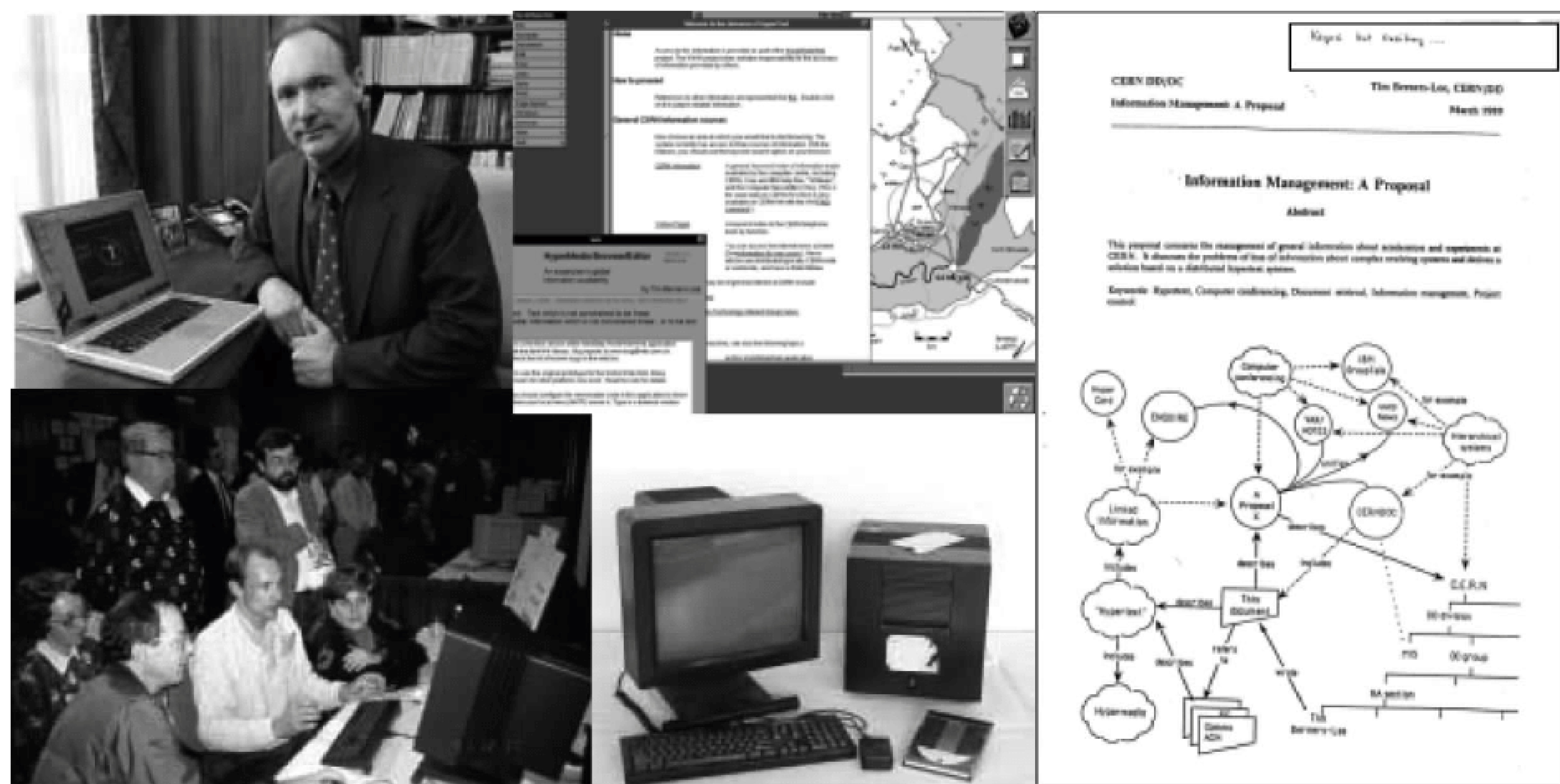


图 1-22　Web 技术发明者 Tim Berners-Lee

2. 浏览器大战

University of Illinois 的 Marc Andreessen 开始开发第一个图形化浏览器 Mosaic，并于 1993 年 2 月发布了第一版的 Mosaic 软件。Mosaic 浏览器很受用户的欢迎，到 1992 年年底已经有 200 台 Web 服务器在运行。1994 年，Marc Andreessen 与 Jim Clark 创办 Netscape 公司，专门开发 Web 客户端与服务器软件。Netscape 公司于 1995 年上市时，投资者普遍认为它是下一个 Microsoft。在 Netscape 公司只有一种产品的情况下，人们花 15 亿美元来购

买他们的股票。1995年,很多大学生每天使用Mosaic与Netscape浏览器在网上冲浪。这时,有些公司已开始在商务事务处理中使用Web技术。1996年,Microsoft公司开发自己的Web浏览器。在以后的三年时间中,Netscape公司的Navigator与Microsoft公司的Internet Explorer进行了一场"浏览器大战"。1998年,America Online以42亿美元收购Netscape公司。1994年,CERN和MIT共同倡议建立了WWW联盟,有几百所大学和公司加入这个联盟。Tim Berners-Lee担任WWW联盟的主管。WWW联盟致力于进一步开发Web技术,以及进行相关协议的标准化工作。

由于Web技术的出现,使Internet从最初主要由计算机专家和大学生科学研究的应用,变为一种广泛使用的信息交互工具。更为难能可贵的是,Tim Berners-Lee并没有为他发明的Web技术申请专利,而是无偿地将它奉献给了全世界所有的人,使我们大家都能够自由地分享这一划时代的技术。Web出现使网站数量和网络通信量呈指数规律增长。Web服务是Internet最方便与受用户欢迎的服务类型,它的影响力也远远超出了专业技术范畴,已广泛应用于电子商务、电子政务、远程教育、远程医疗与信息服务等领域,并且有继续扩大的趋势。

基于Web技术的各种应用的扩展,Internet不仅是一种资源共享、数据通信和信息查询的手段,还逐渐成为人们了解世界、讨论问题、购物休闲,乃至从事学术研究、商贸活动、教育,甚至是政治、军事活动的重要领域。Internet的全球性与开放性,使人们愿意在Internet上发布和获取信息。Web浏览器、超文本标记语言、搜索引擎、Java跨平台编程技术的产生,对Internet的发展产生了重要的作用,使Internet中的信息更丰富,使用更简便。

问题1-6:移动互联网经历了怎样的发展与演变过程?

我们已经清晰地看到:将互联网与移动通信网融为一体的移动互联网技术,已经成为当前技术应用与产业发展的热点。在过去的10年中,人们已经完成了计算机互联网技术与电信移动通信网技术、业务的融合,未来10年将是移动互联网大发展的历史机遇期。

了解从互联网到移动互联网的自然的发展与演变的历程,首先需要了解无线网络技术研究的发展过程。

1. 移动分组网与PRNET、SATNET、ALOHANET

移动分组网的研究可以追溯到20世纪70年代初,其中有两个有代表性的无线分组网是PRNET与ALOHANET。

1) PRNET与SATNET

20世纪70年代初,ARPA资助了分组无线网(PRNET)与分组卫星网(SANET)的研究项目。1977年6月,研究人员第一次实地进行了无线分组跨洋传输的实验。研究人员在加州San Francisco海滨公路一辆行驶的货车上,使用一台LSI-11计算机通过无线分组网PRNET向SRI发送数据,斯坦福研究院SRI再将分组数据通过ARPANET发送到东海岸,然后通过SANET卫星通信网络将分组发送到挪威,从挪威经由海底电缆将分组转发到伦敦;伦敦的计算机再通过SATNET的卫星地面站将分组传回到美国USC的DECKA-10计算机。分组经过往返距离为9.4万英里的传输,没有发生传输错误,证明了TCP的有效性。当时实验系统使用的IP地址的网络地址长度为8b,主机地址长度是24b。图1-23是PRNET、SANET传输路径与PRNET实验车辆、车内网络设备的照片。

PRNET、SANET开启了无线网络技术研究的先河,证明了分组交换技术在无线通信

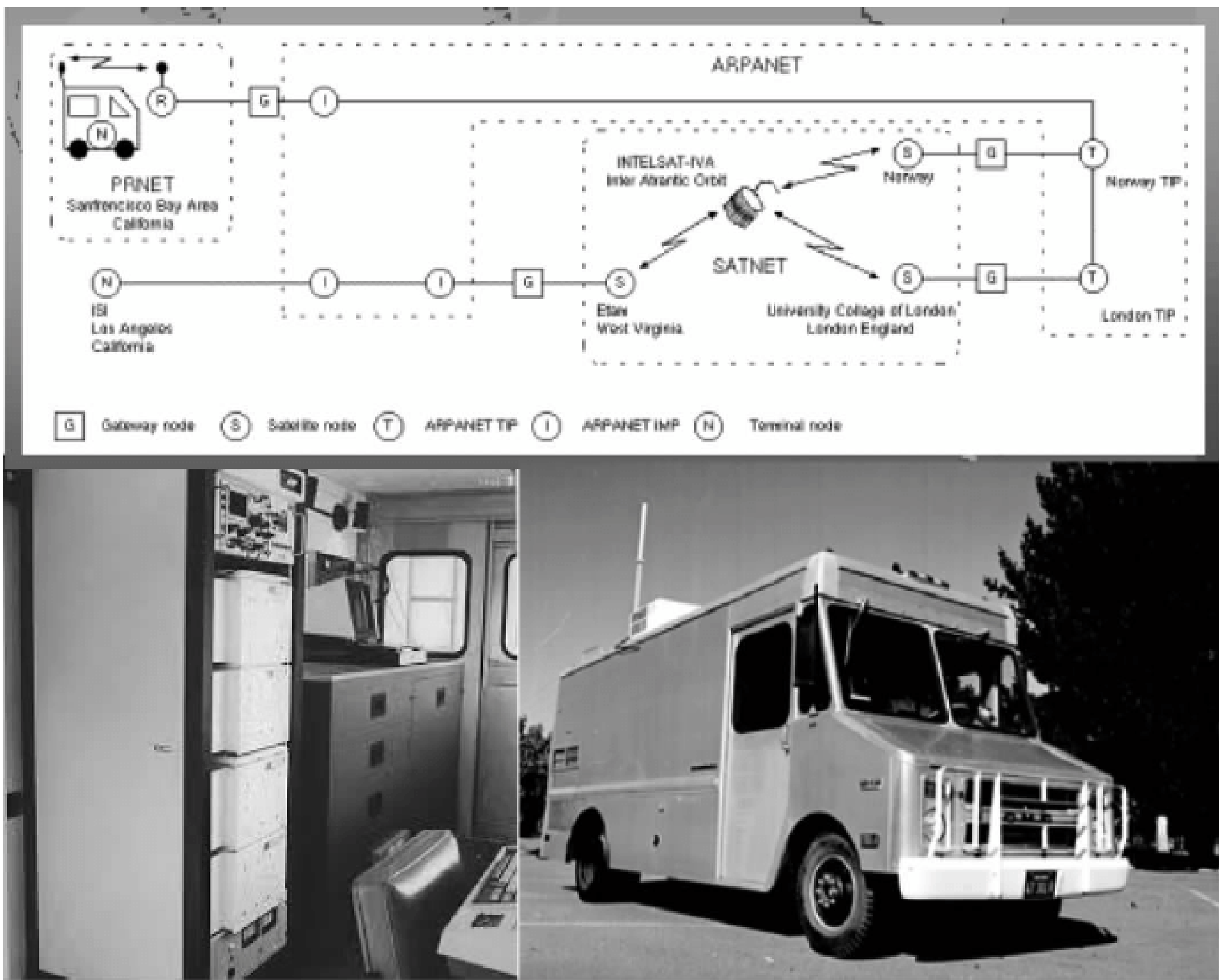


图 1-23　无线分组交换网——PRNET 与 SATNET

领域应用的可行性。在这之后，从事计算机网络与从事移动通信网的电信业技术人员分别从两个方向，开展了更为广泛的无线网络与移动互联网技术与应用的研究工作。

2）ALOHANET

ALOHANET 出现在 20 世纪 60 年代末。夏威夷大学的 Norman Abramson（诺曼·艾布拉姆森）教授和同事们为了在位于夏威夷各个岛屿上的不同校区之间进行计算机通信，研究了一种以无线广播方式工作的分组交换网 ALOHANET。图 1-24 是 Norman Abramson 教授与 ALOHANET 网络设备。



图 1-24　Norman Abramson 教授与 ALOHANET

ALOHANET 使用的是一个共用的无线电信道，支持多个节点对一个共享的无线信道的“多路访问”。ALOHANET 中心节点是一台位于 Oahu 岛校园的 IBM 360 主机，它要通

过学校的无线通信网与分布在各个岛屿校区的计算机终端通信。最初设计时的数据传输速率为4800bps,后来提高到9600bps。ALOHANET的信道方向的规定是以IBM 360主机为基准,从IBM 360主机到终端的无线通信信道为下行信道,从终端到IBM 360主机的无线通信信道为上行信道。下行信道将IBM 360主机的数据分组"广播"到各个岛屿校区的计算机终端,这里不存在冲突问题。但是,当多个不同校区的计算机终端利用上行信道向IBM 360主机传输数据分组时,就可能出现同时有两个或两个以上的终端争用一个通信信道而产生"冲突"的情况。解决"冲突"的办法只有两种:一种是集中控制的方法,另一种是分布控制的方法。集中控制是一种传统的方法,需要在系统中设置一个中心控制节点,由中心控制节点决定哪个终端可以使用共用的上行信道发送数据。但是,由于系统中存在着一个控制中心,因此控制中心会成为系统性能与可靠性瓶颈。ALOHANET执行信道访问控制功能的MAC层采用了分布式控制算法——载波侦听多路访问(Carrier Sense Multiple Access,CSMA)方法。CSMA方法对于之后出现的Ethernet、Wi-Fi,以及射频标签RFID与读写器的通信控制方法的研究都有着重要的指导意义。

2. 无线广域网、无线城域网、无线局域网、无线个人区域网与无线人体区域网研究

随着个人计算机、局域网的日趋广泛,人们逐渐不满足台式计算机与笔记本以固定方式接入互联网的限制,为了满足用户在移动状态下随时随地访问互联网的需求,无线网络与无线接入技术的研究与应用成为网络的热点课题,无线广域网(Wireless Wide Area Network,WWAN)、无线城域网(Wireless Metropolitan Area Network,WMAN)、无线局域网(Wireless Local Area Network,WLAN)与无线个人区域网(Wireless Personal Area Network,WPAN)、无线人体区域网(Wireless Bady Area Network,WBAN)技术与标准逐渐成熟并进入实用阶段。

1) 无线网络的特征与分类

无线网络的特征与分类如图1-25所示。

无线网络的特征与分类可以从4个方面来认识。

(1) 传输方式。

传输方式涉及无线网络采用的载波类型、频段与调制方式。目前无线网络主要采用无线频段与红外频段。调制方式又可以进一步分为扩展频谱、窄带调制与红外方式。

采用扩展频谱方式的无线网络一般选择ISM频段,如跳频扩频(FHSS)与直接序列扩频(DSSS)。在窄带调制方式中,数据基带信号的频谱不做任何扩展,而是将基带信号的频谱直接调制在载波上发射出去。由于红外传输技术相对传输,不易受干扰,因此在近年得到很大的发展,目前广泛使用的家电控制几乎全部使用红外方式传输。在以上三种传输方式中,扩展频谱与红外方式不需要申请频段,而窄带调制方式需要向国家无线电管理委员会申请频段。

(2) 网络拓扑。

无线网络拓扑基本上分为两类:需要建立基站的有中心无线网络,不需要建立基站、采用对等方式(Peer-to-Peer)通信的无线网络。例如,在IEEE 802.11标准中,WLAN方式需要设置基站AP,移动终端采用"一跳"与"竞争"的方式通过基站接入网络;而Ad Hoc方式采用"多跳"的方式,以自组网的方式构成移动的无线网络。



图 1-25　无线网络的特征与分类

(3) 网络接口。

无线网络中节点网络可以选择在物理层,或者在数据链路层(即 MAC 层)。物理层接入实际上是用无线信道代替了有线网络中的有线信道,物理层以上的各层基本保持不变。MAC 接入不采用传统局域网的 MAC 层协议,而是采用更能够适合无线信道的 MAC 层访问控制协议。网络层及以上的协议基本保持不变。

(4) 无线网络分类。

根据网络覆盖范围的不同,无线网络可以分为:

① 无线广域网(WWAN);

② 无线城域网(WMAN);

③ 无线局域网(WLAN);

④ 无线个人区域网(WPAN);

⑤ 无线人体区域网(WBAN)。

无线网络分类相对应的协议标准如图 1-26 所示。

图 1-27 描述了各种无线网络覆盖范围、带宽与主要的协议标准。

2) 无线广域网

无线广域网主要包括两类技术:卫星通信与移动通信。

从计算机网络发展历史的讨论中已经看出,20 世纪 70 年代初与 ARPANET 同期研究的无线分组网 PRNET 与分组卫星网络(SANET),实际上就是最早出现的无线广域网。

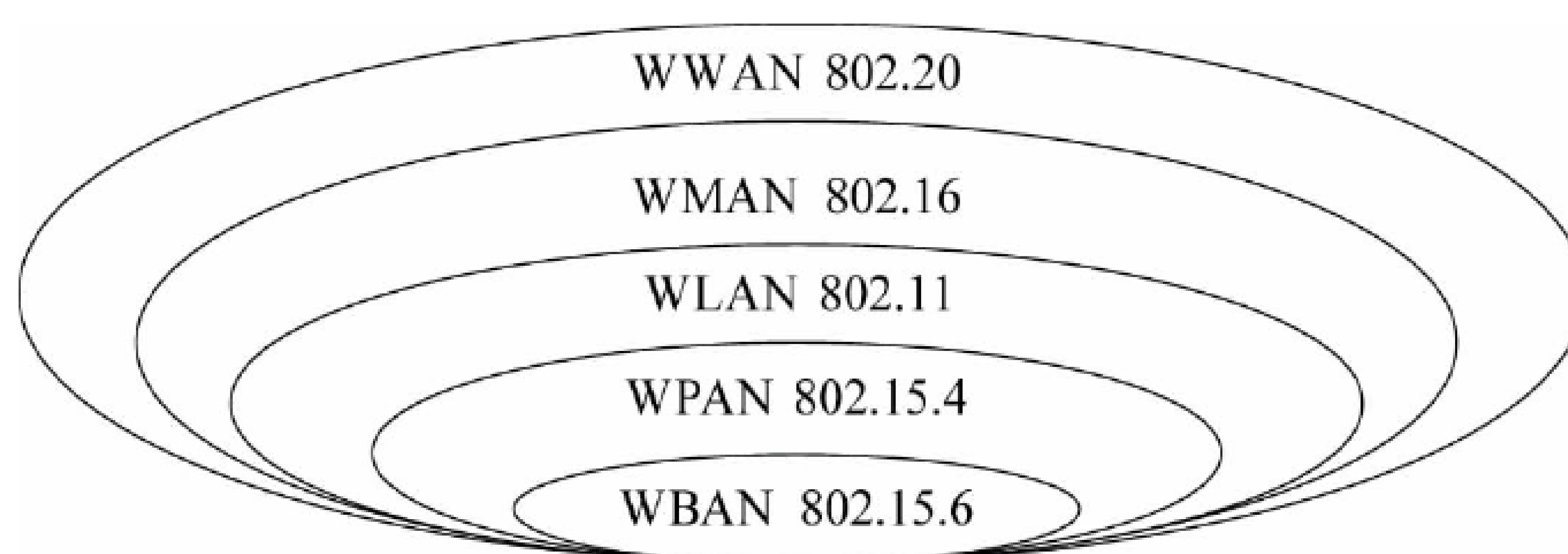


图 1-26　无线网络分类与标准

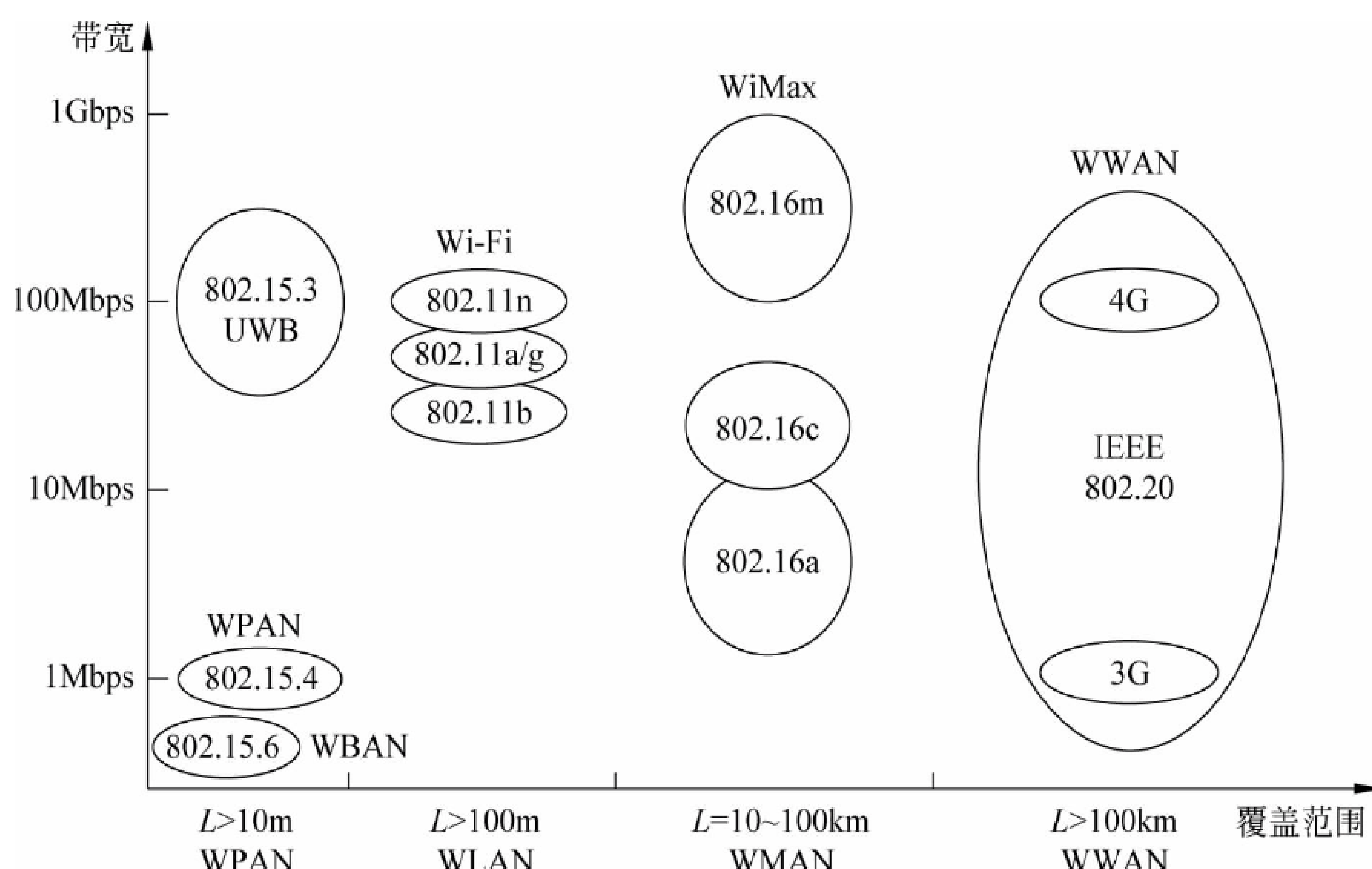


图 1-27　无线网络技术的比较

由于卫星通信具有通信距离远、费用与通信距离无关、覆盖面积大、不受地理条件的限制、通信信道带宽宽、可进行多址通信与移动通信的优点,因此它在最近的三十多年里获得了迅速的发展,并成为现代主要的通信手段之一。

商用通信卫星一般被发射在赤道上方 36 000km 的同步轨道上。这就意味着,当地球自转时,同步卫星也以一个适当的速度沿地球自转方向绕轨道运行,地球与卫星之间可以保持相对的静止。三颗这样的卫星均匀沿轨道分开,就可以覆盖整个地球表面。

20 世纪 90 年代初,由于小卫星技术的飞速发展,出现了中、低轨道卫星移动通信的新方法。中、低轨道卫星不是同步卫星,它作为陆地移动通信系统的补充和扩展,与地面公用通信网有机地结合起来,才能实现全球个人移动通信。低轨道移动通信系统实现个人通信的优点是:卫星轨道高度低,使传输延迟缩短,多个卫星组成的星座可以真正覆盖全球。卫星移动通信系统将形成一个空间的通信子网,它将实现物理层、数据链路层与网络层的功能。当时提出的低轨道卫星方案的大公司主要有 8 家,其中最有代表性的低轨道卫星移动通信系统主要有铱(Iridium)系统、Globalstar 系统、Arics 系统、Leo-Set 系统、Coscon 系统、Teledesic 系统等。

铱系统是美国 Motorola 公司提出的一种利用低轨道卫星群,实现全球卫星移动通信的建设计划,耗资约 34 亿美元。铱系统由 66 颗小型智能卫星,均匀有序地分布在离地面

785km 上空的 6 个轨道平面上。每颗星可以提供 48 个点波束、每个波束平均包含 80 个信道，共 3840 个全双工电路信道。每颗星投射的多波束在地球表面上形成 48 个蜂窝区，每个蜂窝区的直径约为 667km，总覆盖直径约 4000km，全球共有 2150 个蜂窝系统。同时，智能卫星可以利用星与星之间的通信，实现空间数据交换与路由的功能。图 1-28 给出了铱系统示意图。

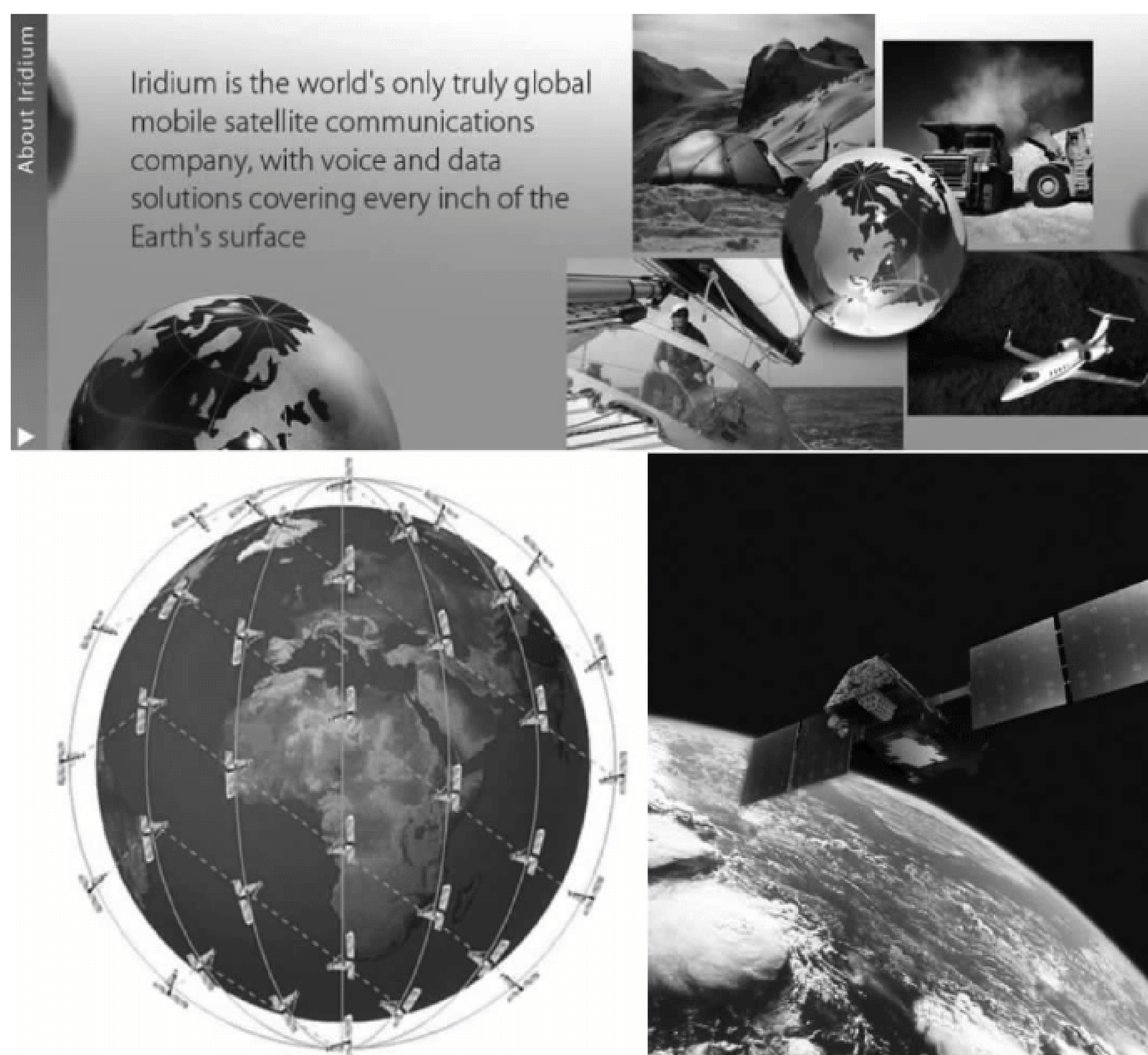


图 1-28　典型的全球低轨道卫星通信网——铱系统示意图

1990 年 6 月，铱系统建设计划首次宣布，1992 年 9 月得到美国 FCC 的许可证。铱通信卫星于 1998 年 11 月 1 日开始提供通信服务。但是，由于受到地面移动电话（如 GSM 系统）的挤压，加上自身昂贵的通信费用，使得铱系统使用率不高，造成铱卫星计划执行中的财务困难。铱系统建设计划最终于 1999 年 8 月 13 日宣布失败。不过部分在轨的铱通信卫星仍然在为美国军方提供通信服务。尽管铱系统建设计划宣布失败，但是它对目前开展的星际网络的研究工作起到了重要的指导作用。

蜂窝移动通信网的设计涉及 OSI 参考模型的物理层、数据链路层与网络层。3G/4G/5G 移动通信系统充分利用了地面移动通信网、卫星通信网和光纤通信网的互联，组成了一个全球无缝覆盖的移动通信网络，也为构成无线广域网提供了重要的技术支持。

IEEE 802.20 是无线广域网 WWAN 的重要标准。IEEE 802.20 标准的研究首先是由 IEEE 802.16 工作组在 2002 年 3 月提出，并为此于 2002 年 9 月成立了 IEEE 802.20 工作组。IEEE 802.20 标准是为了有效解决无线广域网中移动性与传输速率相互矛盾的问题，研究了一种适用于高速移动环境下的宽带无线接入系统空中接口规范。

IEEE 802.20 协议模型覆盖 OSI 参考模型的物理层与数据链路层。IEEE 802.20 标准在设计理念上采用基于分组数据的纯 IP 结构，适应互联网突发性数据业务，其性能优于 3G/3.5G 技术。但是，从技术角度看 IEEE 802.20 是对 3G 和 IEEE 802.16e 标准的一个补充。至于 IEEE 802.20 标准是否能够取代二者的地位，前景并不明朗。从经济可行性角度

考虑,目前移动运营商已经在移动通信系统3G/4G中投入巨资购买牌照、部署网络,不可能放弃现有投资而重新部署新的网络。从目前实际情况看,IEEE 802.20技术标准本身仍有待完善,产品市场与产业链没有形成,所以还很难判定它在未来移动广域通信网市场中的位置。这是造成IEEE 802.20标准与技术发展迟迟不能发展的主要原因。

3）无线城域网

由于在城市的一些大楼和分散的社区,铺设电缆与光纤的费用往往高于无线通信设施建设的费用,因此人们就开始研究如何在市区范围的高楼之间,利用无线通信手段解决局域网与局域网,固定或移动的个人用户计算机接入互联网的问题。1999年7月,IEEE 802委员会成立一个工作组,专门研究宽带无线城域网WMAN标准问题。2002年,该工作组发布了IEEE 802.16宽带无线城域网标准。由业界成员参加的WiMax论坛致力于IEEE 802.16标准的推广与应用。因此人们通常用“WiMax”表示按照IEEE 802.16标准组建的无线城域网。图1-29给出了WiMax概念与结构示意图。

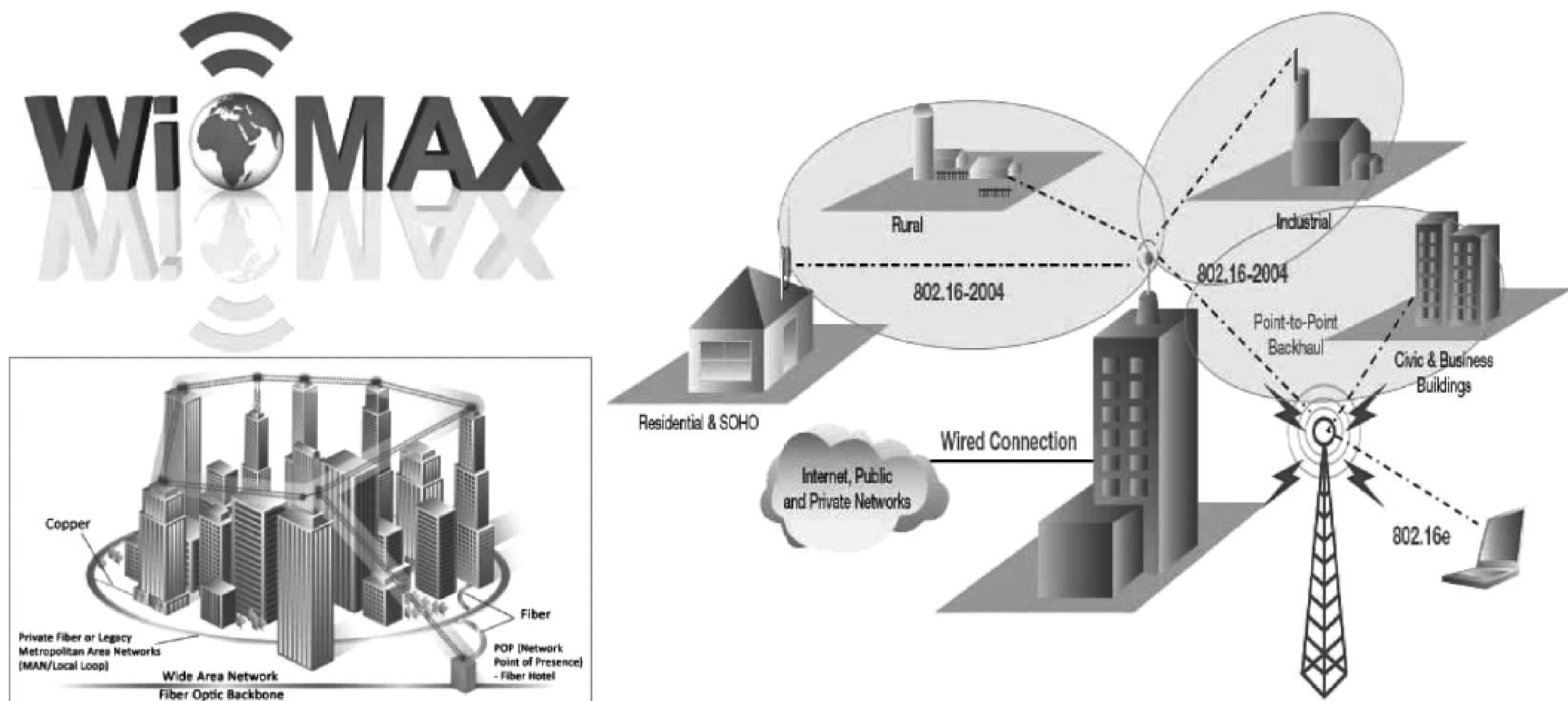


图1-29　WiMax概念示意图

按IEEE 802.16标准建设的无线网络需要在每个建筑物上建立基站。基站之间采用全双工、宽带通信方式工作,以满足固定节点以及火车、汽车等移动物体的无线通信需求。2011年4月,IEEE通过了802.16m标准,该标准是为下一代无线城域网而设计。IEEE 802.16m标准可在固定的基站之间提供1Gbps的数据传输速率,为移动用户提供100Mbps的数据传输速率。

4）无线局域网

我们在学校、机场、车站、咖啡厅等很多场合随处可见“Wi-Fi”标记,这也从一个方面说明了Wi-Fi技术的重要性。

无线局域网以微波、激光与红外等无线载波作为传输介质,代替传统局域网中的同轴电缆、双绞线与光纤,实现移动节点的物理层与数据链路层功能。无线局域网(WLAN)是实现移动计算网络的关键技术之一,技术与标准发展速度相当快。IEEE 802委员会成立了802.11工作组专门从事无线局域网的研究,并于1997年公布了IEEE 802.11无线局域网标准。从1999年到2006年是无线局域网发展最快的阶段,有几十项IEEE 802.11协议标准(如IEEE 802.11a、802.11b、802.11n等)颁布,它们涉及不同无线频段、不同速率、不同

组网方式与不同安全认证方式的无线局域网技术。

无线局域网可以作为传统局域网的扩充，也可以用于漫游访问、建筑物之间的互连。图 1-30 给出了无线城市、Wi-Fi 接入点标志与应用情景的示意图。目前，很多大学的校园网都在一些没有预先埋设局域网接口的教室、图书馆与自习室，以及校园内的草坪中安装了无线局域网接入设备——接入点(AP)。学生可以自由地在教室、图书馆与自习室，校园中任意看书的地方，随时随地使用笔记本、智能手机、iPad、PDA 查阅校园网上教学文档，检索图书馆馆藏的文献资料，提交作业，发送和接收电子邮件，访问 Web 网站。目前，Wi-Fi 已经广泛应用于办公楼、家庭、咖啡厅、机场候机厅，甚至是火车、飞机上。



图 1-30　无线城市、Wi-Fi 接入点标志与应用情景

5）无线个人区域网

无线个人区域网技术、标准与应用是当前网络技术研究的热点之一。尽管 IEEE 希望将 802.15.4 推荐为近距离范围内移动办公设备之间的低速互连标准，但是业界已经存在两个有影响力的无线个人区域网技术与协议，即蓝牙技术与 ZigBee 技术。

IEEE 在 2000 年正式成立了 802.15 工作组，致力于低速无线个人区域网通信标准的研究与制定工作。LR-WPAN 研究的目标是解决近距离、低速率、低功耗、低成本、低复杂度的嵌入式无线传感器，以及自动控制设备、自动读表设备之间的数据传输问题。按照计算机网络体系结构的设计思想，IEEE 802.15.4 协议只包括物理层与数据链路层。

目前，无线传感器网络 WSN 研究的平台大多数都采用 IEEE 802.15.4 标准(如图 1-31 所示)。

1994 年，Ericsson 公司看好移动电话与无线耳机的连接，以及笔记本与鼠标、键盘、打印机、投影仪的无线连接技术与市场，对于近距离的无线连接产生了浓厚的兴趣。Ericsson 公司与 IBM、Intel、Nokia 和 Toshiba 等 4 家公司，发起开发一个短距离、低功耗、低成本通信标准和技术的倡议，并将它命名为“蓝牙(Bluetooth)”无线通信技术。蓝牙可以解决各种智能设备，例如笔记本与键盘、鼠标，以及智能手机、PDA、数码相机、摄像机、耳机之间的无线通信问题。

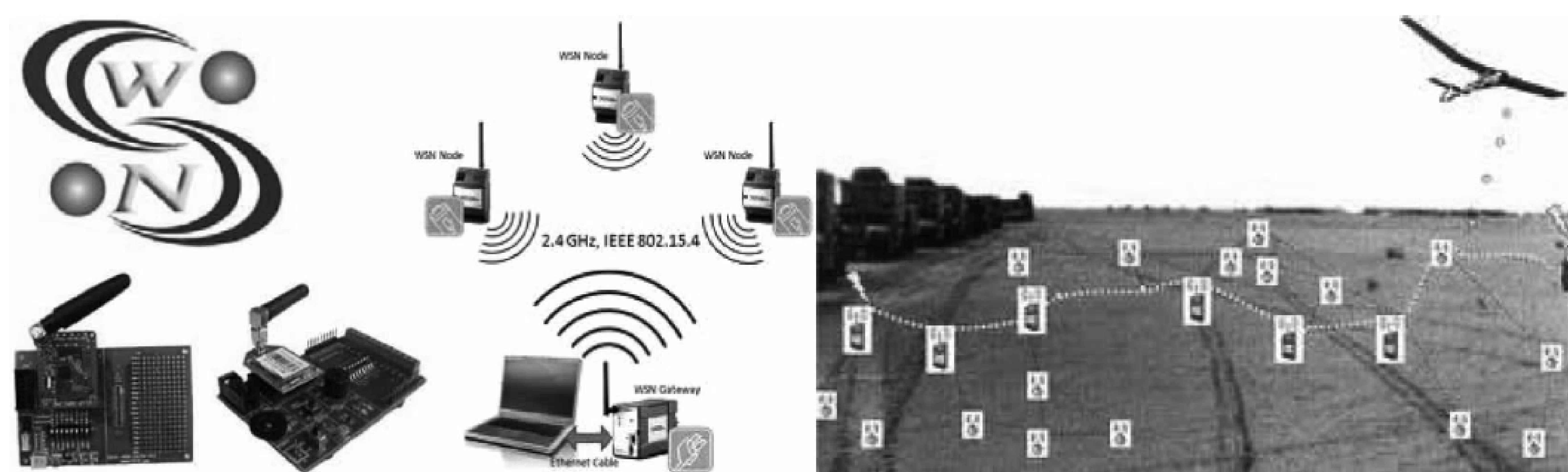


图 1-31　采用 IEEE 802.15.4 标准研发的 WSN 示意图

蓝牙通信采用不需要专门申请的工业、科学与医学(ISM)频段。工作频率在 2.4GHz 时,数据传输速率最高为 1Mbps,通信距离一般在 10cm～10m,支持点对点、点对多点的通信。目前,采用蓝牙技术开发的键盘、鼠标、耳机、MP3 播放器、投影仪(笔)、音箱已经广泛使用(如图 1-32 所示)。



图 1-32　蓝牙技术的应用

大多数无线网络通信协议不涉及应用层的问题。IEEE 802.3 与 802.11 是按计算机网络体系结构的思想而设计,它只回答物理层与数据链路层的问题,并不涉及高层协议。蓝牙技术的设计思路不一样。蓝牙规范 1.0 版规定 13 种应用所需的专门协议集。由于蓝牙系统的工作范围不大,从网络层到传输层都必须设计得很简单,有可能在一个通信协议的设计中考虑具体支持某种应用。但是,这种做法可能导致协议的庞大与复杂,因此蓝牙规范 1.0 版长达 1500 页就可以理解。

1998 年 5 月,由 Ericsson、Intel、IBM、Nokia 和 Toshiba 等公司发起蓝牙技术联盟

(SIG)。目前,SIG 共有一千八百多个成员,包括消费类电子产品制造商、芯片制造厂家与电信业等。SIG 的主要任务致力于蓝牙技术的推广。目前,蓝牙技术已出现了很多版本,传输速率高达 480Mbps、传输距离可达到几十米。

ZigBee 是一种面向自动控制的低速、低功耗、低价格的无线网络技术,目前已经有一些物联网系统开始应用 ZigBee 通信技术。

ZigBee 协议标准的第一版是 2004 年完成,第二版是 2006 年颁布。ZigBee 的通信速率要求低于蓝牙,要求由电池供电,在不更换电池的情况下,能工作几个月甚至几年。同时,ZigBee 网络的节点数量、覆盖规模比由蓝牙技术支持的网络大得多。ZigBee 无线设备工作在公共频道,在 2.4GHz 时传输速率为 250kbps,在 915MHz 时为 40kbps。ZigBee 的传输距离为 10~75m。

由芯片制造商、OEM 厂商、应用系统开发商、与无线传感器网络开发商共同成立 ZigBee 联盟。ZigBee 联盟是一个国际性的非营利工业技术团体,它的任务是开发和推广 ZigBee 标准与技术。ZigBee 应用领域集中在 5 个方面:家庭网络、智能能源、建筑自动化、远程通信服务与个人健康助理等。ZigBee 适应于数据采集与控制的节点多、数据传输量不大、覆盖面广、造价低的应用领域,在家庭网络、安全监控、医疗保健、工业控制、无线定位、无线读表、智能玩具、智能农业等方面展现重要的应用前景。图 1-33 给出了采用 ZigBee 技术开发的人体健康状况监测系统结构示意图。



图 1-33 采用 ZigBee 技术研发的应用系统示意图

6) 无线人体区域网

作为近距离无线通信,虽然已经存在个人区域网 PAN 的概念,但是针对医疗及保健的应用,仅限人体周边更短传输距离的应用有其特殊性。随着物联网在医疗健康、疾病监控和预防中的应用越来越广泛,研究可穿戴设备与植入人体内的生物传感器组成的无线人体传感器网络(Wireless Body Sensor Network,WBSN)成为新的无线传感器网络新的研究热点。无线人体传感器网络也称作生物医疗传感器网络(Biomedical Sensor Network,BSN)或无线人体区域传感器网络(Wireless Body Area Sensor Network,WBASN)。无线人体区域传感器网络也经常缩写为体域网(Body Area Network,BAN)或无线人体区域网(WBAN)。IEEE 于 2012 年正式批准了无线体域网 WBN 标准——IEEE 802.15.6。本章统一使用无线人体传感器网络(WBSN)。

WBSN 的研究希望为健康医疗监控应用提供一个集成硬件、软件的无线通信平台,特别强调适应于可穿戴与可植入的生物传感器的尺寸,以及低功耗的无线通信的要求。

2007 年开始,IEEE 的 802.15 工作组 TG6 开始了无线人体区域网络及通信标准的研究,经过约五年时间终于完成了标准制定工作。IEEE 802.15.6 标准研究传输速率最高为 10Mbps、最长传输距离为 1m 的无线传输技术,以取代蓝牙与 ZigBee 等标准。

IEEE 802.15.6 除了应用于医疗保健与疾病控制之外,也可以用于日常生活中便携播放器与无线耳机之间等人体身边便携式装置之间的通信,以及消防、探险、军事等特殊场合的应用。图 1-34 给出了无线人体传感器网络基本概念与应用场景示意图。

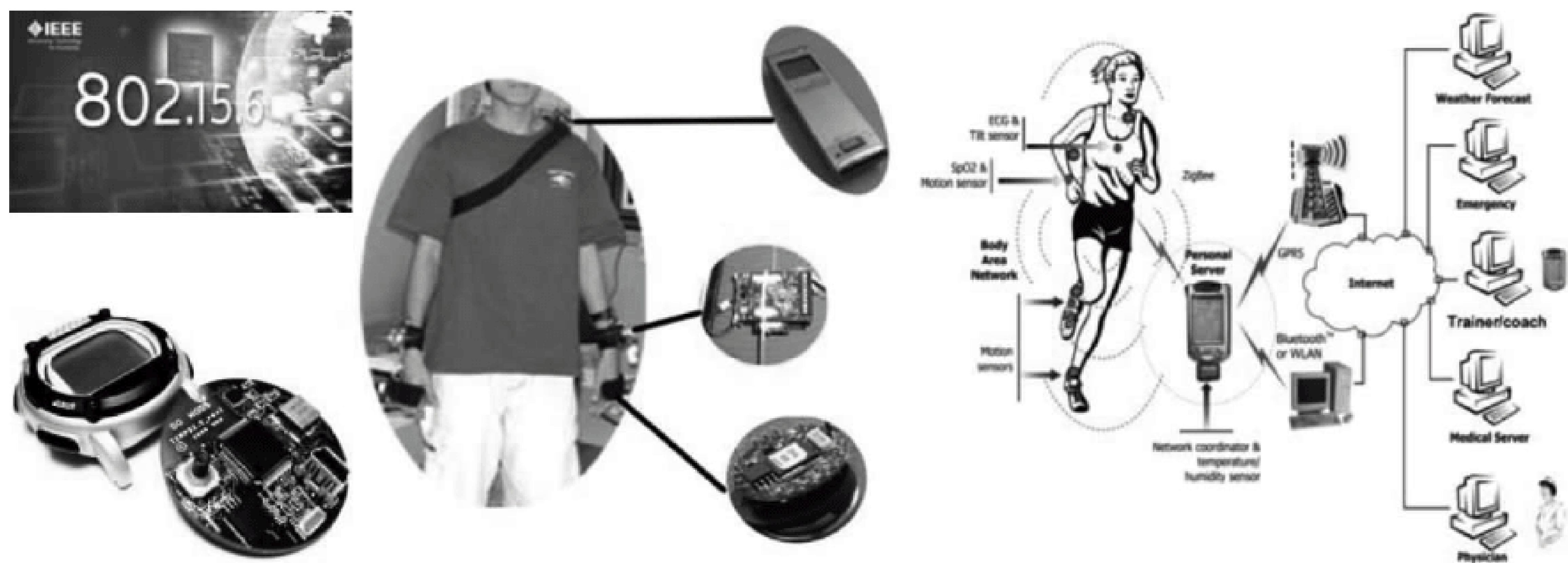


图 1-34　WBSN 概念与应用场景示意图

目前的研究人员在 IEEE 802.15.6 基础上开展的研究工作主要集中在:WBSN 中情景感知和周围环境感知;WBSN 可穿戴性、可扩展性和资源优化;基于多种通信方式构建混合的 WBSN;移动 WBSN 中跟踪和能量感知 MAC 算法;从低能耗和通信的角度构建新型的 WBSN 系统的 WBSN 的架构;WBSN 中的数据融合技术;WBSN 对人体活动的监控;WBSN 的自适应性和可调节性、中间件、信号处理算法、健康及活动监控和网络可靠性。

问题 1-7:如何认识物联网发展的技术背景与社会背景?

我们知道,任何一项重大科学技术发展的背后,都必然有其深厚的社会发展背景与技术发展背景。在分析物联网发展的社会背景时,人们总会谈到 4 件事,那就是:比尔·盖茨的《未来之路》、美国 MIT Auto-ID 实验室开展的基于射频标签 RFID 的产品电子代码 EPC 技术的研究、IBM 公司提出的智慧地球的概念,以及国际电信联盟 ITU 发布的互联网研究报告 *The Internet of Things*。

1. 比尔·盖茨与《未来之路》

首先来看看,比尔·盖茨在他的《未来之路》一书中描述他对未来互联网时代的憧憬,与现在讨论的物联网概念的产生究竟有什么关系。

大家知道,20 世纪 90 年代互联网应用与信息高速公路建设开始进入高速发展阶段,社会上赞成的意见与反对的意见争论很激烈。1995 年,比尔·盖茨出版了《未来之路》一书。他在前言里写到:他写这本书的目的就是要向世人介绍未来的互联网时代将会发生哪些变化。他希望通过这本书,描述他对未来互联网时代的憧憬,同时希望起到"促进理解、思考"的作用。

1995 年,比尔·盖茨出版了《未来之路》一书。书中第 1 章的名字叫作"一场革命开始了",比尔·盖茨在这一章中提出了 10 个问题。第 6 个问题是"位于哪里的哪一家商店在明

天早晨以最低价把一个测量你脉搏的手表送货上门?"

在这之后,比尔·盖茨设想了一个场景:假定你在开车的时候想找一家新餐馆,并想看看它的菜单、红酒单和当天的特色菜。计算机系统可以帮你找到。那么你需要预订座位、一张地图,了解目前的交通情况。当你发出指令之后,可以一边开车,一边等待计算机系统打印,或者是通过语音方式朗读出来给你听,并且你获得的信息是实时和不断更新的。现在再次读到这一段文字时,我们不能不联想到当前谈论的移动互联网的"基于位置的服务"与物联网的"智能交通"的应用场景。

在《未来之路》的第10章"不出户,知天下"中,比尔·盖茨用两句话来描述他在西雅图华盛顿湖畔的住所,他说"我的房子用木材、玻璃、水泥、石头建成",同时"我的房子也是用芯片和软件建成的"。

《未来之路》的第10章"不出户,知天下"中提出了"物-物互联"的设想。比尔·盖茨在书中引入了一种"电子别针"。当客人进入住所时,第一件事是别上一个电子别针,这根电子别针把来访者与房子里面的各种电子服务系统"连接"起来了。凭着来访者戴的电子别针,房子的自动感知系统知道"你是谁""你在哪里""需要为你提供什么样的服务",房子将通过这些信息尽量满足、甚至是预见你的需求。当你沿着大厅走时,你前面的灯光会逐渐变强,而后面的灯光正在消失。音乐会随着来访者一起移动,而其他人却听不到声音。电影与新闻将能够跟着你在房子里移动。如果有一个需要来访者接的电话,那么只有离来访者最近的电话机才会响。手持遥控器能够扩大电子别针的控制能力。来访者可以通过遥控器发出指令,能够从数千张图片、录音、电影、电视节目中选择你所需要的信息。

所有这些描述与我们目前对于物联网中物-物相连的设想非常相似,这也证实了比尔·盖茨的一个预言:"我要用的技术在现在是实验性的,但过一段时间我正在做的部分事情会被广为接受。"他在描述自己住所的未来发展前景时说:"微处理器芯片和存储器的安装,以及控制它们运行的软件,这些都会在最近几年里随着信息高速公路进入数百万个家庭。"现在读起来这些话,我们会发现这与我们现在描述的物联网实现"物理世界与信息世界的融合""智能家居"应用的思路是如此吻合。

2. Auto-ID实验室、RFID与物联网的概念

1998年,美国麻省理工学院Auto-ID实验室的研究人员在成功地完成了产品电子代码(Electronic Product Code,EPC)研究的基础上,提出了利用射频标签(Radio Frequency Identification,RFID)、无线网络与互联网,构建物-物互联的物联网的概念与解决方案。

RFID技术能够满足物品信息自动、快速、准确识别的需求。当RFID技术与互联网技术结合在一起时,就可以构建覆盖全世界的物联网智能物流应用系统。

基于EPC的物联网应用系统是建立在互联网的基础之上的,同时它也需要增加必要的物联网基础设施,那就是:对象名字服务(ONS)机制、对象名字服务器与服务器体系,以及EPC信息服务机制、EPC信息服务器与服务器体系。

EPC研究工作对物联网技术发展的影响,主要表现在以下三个方面。

(1) EPC研究向我们展现出一种基于RFID的物联网智能物流应用原型系统。

(2) EPC研究向我们展现出"物-物互联"的物联网的概念与解决方案。

(3) EPC研究向我们展现出物联网与互联网相互促进、共同发展的协同工作关系。

3. 物联网概念与ITU的研究报告

“物联网”概念产生于20世纪90年代,而真正引起各国政府与产业界的重视是在2005年国际电信联盟(ITU)的“信息社会世界峰会(WSIS)”之后。2005年度报告的题目是“物联网(Internet of Things)”。

在2005年突尼斯举行的信息社会世界峰会上,ITU发布的报告系统地介绍了意大利、日本、韩国与新加坡等国家的案例,并提出了“物联网时代”的构想。报告预见:RFID、传感器技术、智能嵌入式技术及纳米技术将广泛应用。

4. 物联网与智慧地球

国际金融危机爆发以来,为了尽快摆脱危机的影响,很多国家都在寻求和培育新的经济增长点。2009年1月28日,美国总统奥巴马在与美国工商界领袖举行的圆桌会议上听取了IBM公司首席执行官彭明盛关于“智慧地球”的报告。

IBM公司在智慧地球概念的基础上提出了他们对物联网的理解。IBM的学者认为:智慧地球将传感器嵌入和装备到电网、铁路、桥梁、隧道、公路、建筑、供水系统、大坝、油气管道等各种物体中,并通过超级计算机和云计算组成物联网,实现人与物的融合。智慧地球的概念是希望通过在基础设施和制造业中大量嵌入传感器,捕捉运行过程中的各种信息,然后通过无线网络接入到互联网,通过计算机分析、处理和发出指令,反馈给控制器,远程执行指令。控制的对象小到一个开关、一个可编程控制器、一台发电机,大到一个行业。通过智慧地球技术的实施,人类可以以更加精细和动态的方式管理生产与生活,提高资源利用率和生产能力,改善人与自然的关系。彭明盛建议政府投资新一代智慧型基础设施的建设,以此拉动美国经济的增长,渡过经济危机的难关。2009年1月7日,IBM公司与美国智库信息技术与创新基金会(ITIF)共同向美国政府提交了名为“*The Digital Road to Recover: A Stimulus Plan to Create Jobs, Boost Productivity and Revitalize America*”的建议书。建议书提出通过信息通信技术(ICT)投资,可以在短时间创造就业机会。美国政府在智能电网、智能医疗与宽带网络这三个领域新增投资300亿美元,可以为美国民众创造出94.9万个就业机会。

美国总统奥巴马明确表示:“经济刺激资金将会投入到宽带网络等新兴技术之中,毫无疑问,这就是美国在21世纪保持和夺回竞争优势的方法。”美国政府将“宽带网络等新兴技术”定位为振兴经济、确立美国全球竞争优势的关键战略,随后出台了总额为7870亿美元的《经济复苏和再投资法》,以落实上述计划。美国国家情报委员会(NIC)发表的“2025年对美国利益有潜在影响的关键技术”报告中,将物联网列为6大关键技术之一。物联网与新能源成为美国摆脱经济危机、振兴经济的两大核心武器。

IBM前首席执行官郭士纳曾经提出过一个重要的观点——“十五年周期定律”。他认为,计算模式每隔15年发生一次变革。按照他总结出来的规律,1965年前后出现了以大型计算机为标志的变革,1980年前后出现了以个人计算机普及为标志的变革,1995年前后出现了以互联网应用为标志的变革,那么2010年前后出现的以物联网为标志的变革将进一步验证了他的预测。物联网作为一种新的计算模式将会引起各国产业结构的变化,甚至会造成国家之间竞争格局的动荡与变化。各国将物联网作为振兴经济、调整产业结构、确立竞争优势的重大战略决策,它将对各国的经济与社会发展产生重大的影响。

问题 1-8：计算机网络存在着几种定义？

从作者近二十多年的教学过程中，能够收集到对计算机网络的定义基本上是三种，即广义观点的定义、资源共享观点的定义，以及用户透明性观点的定义。

人们对计算机网络提出了不同的定义，它们出现在计算机网络发展过程的不同阶段。不同的定义反映着当时网络技术发展的水平，以及人们对网络技术的认识程度。这些定义可以分为以下三类。

1. 广义观点的定义

广义的观点产生于计算机网络发展的第一阶段向第二阶段过渡时期，比资源共享观点的定义提出得早。远程联机系统的发展为计算机应用开辟了新的领域。随着计算机应用的发展，一个大公司或一个部门常常会拥有多台计算机系统，而且这些计算机系统分散在不同的地点，它们之间要经常进行业务信息交换。通过通信网络各地区子公司的计算机可以将不同地点各个子公司的数据汇集后传送到总公司计算机。广义的观点描述了这种以传输信息为主要目的、用通信线路将多个计算机连接起来的计算机系统的集合，我们将它定义为计算机通信网。计算机通信网在物理结构上具有了计算机网络的雏形，但它以相互间的数据传输为主要目的，资源共享能力弱，是计算机网络的低级阶段。

2. 资源共享观点的定义

资源共享观点将计算机网络定义为“以能够相互共享资源的方式互连起来的自治计算机系统的集合”。资源共享观点的定义符合目前计算机网络的基本特征。

3. 用户透明性观点的定义

分布式系统(Distributed System)与计算机网络是两个容易被混淆的概念。用户透明性观点定义计算机网络“存在着一个能为用户自动管理资源的网络操作系统，由它调用完成用户任务所需要的资源，而整个网络像一个大的计算机系统一样对用户是透明的。”严格地说，用户透明性观点的定义描述了一个分布式系统。

问题 1-9：计算机网络与分布式计算机系统有哪些区别？

分布式计算机系统有以下 4 个主要的特征。

(1) 系统拥有多种通用的物理和逻辑资源，可以动态地给它们分配任务。

(2) 系统中分散的物理和逻辑资源通过计算机网络实现信息交换。

(3) 系统存在一个以全局方式管理系统资源的分布式操作系统。

(4) 系统内部结构对用户是完全透明的。

组建一个计算机网络需要有网络硬件与网络系统软件，我们把网络系统软件称作网络操作系统。目前计算机网络操作系统要求网络用户在使用网络资源时必须了解网络资源分布情况。在共享某一台计算机资源时，首先要在这台计算机上登录，在成为该计算机的合法用户后，才能进行允许的资源共享操作。而分布式操作系统以全局方式管理系统资源，自动为用户任务调度网络资源。分布式系统的用户不必关心网络环境中资源的分布情况，以及联网计算机的差异，用户的作业管理与文件管理过程对用户是透明的。计算机网络是一种松耦合系统，而分布式系统是一种紧耦合系统。分布式系统与计算机网络的区别主要不在于它们的物理结构，而是在高层软件。计算机网络为分布式系统研究提供了技术基础，而分

布式系统则是计算机网络技术发展更高级的形式。

分布式系统与计算机网络的区别主要不在于它们的物理结构,而是在高层软件。计算机网络不具备以全局方式管理和调度网络资源的分布式操作系统,它为分布式系统提供了通信平台;而分布式系统在计算机网络的基础上,通过分布式操作系统来管理和调度网络中的计算与存储资源。Web就是一个成功的分布式系统的实例。在Web系统中,用户所有的一切都看起来像在一个Web页面一样。

问题1-10:如何理解计算机网络定义中"独立计算机系统"的含义?

我们可以将计算机网络定义为:以相互共享资源的方式互连起来的自治计算机系统的集合。因此,计算机网络的主要特征是:资源共享、独立计算机系统与遵循共同的网络协议。

有一种观点是:现在的计算机网络连接的不一定是"独立"的计算机系统。这一点需要深入到目前大量联网设备的特征方面来认识。随着Internet与三网融合技术的发展,联网计算机的概念在发生变化。联网计算机的类型已经从大型计算机、个人计算机、个人事务助理,逐步扩展到智能手机、Pad、电视机、家用电器等各种移动或固定的数字终端设备。但是,无论接入网络的数字终端设备类型如何变化,这些接入设备都具有一个相同的特点,那就是:内部都有CPU、操作系统与执行网络协议的软件,都属于端系统中的设备。不同之处是:由于应用领域与功能的不同,接入设备使用的CPU、操作系统与网络软件的性能、规模与功能可能不同。例如,智能手机有自己的嵌入式操作系统,如果不连入Internet,它除了可以打电话之外,还可以作为计算器、词典、记事本等计算功能。很多RFID或传感器结点,它们都是通过一个管理结点接入到计算机网络的。

综上所述,在计算机网络技术的讨论中,我们已经将早期"独立计算机系统"的概念从计算机系统扩展到能够连接到互联网、移动互联网、物联网的各种"计算设备",如智能手机、PDA、Pad、可穿戴技术设备、智能机器人、智能家电、传感器、RFID、智能汽车等。因此,我们在计算机网络技术的讨论中将这些联网的计算设备统称为主机(Host)或结点(Node)。

问题1-11:在计算机网络与互联网结构的研究中采用了几种抽象方法?

计算机网络的抽象方法随着互联网结构不断扩大而演变和发展着。

资源子网与通信子网的两级结构的抽象方法出现在ARPANET研究阶段,但是它实际上是以广域网为对象提出的计算机网络结构的抽象描述方法。该方法能够很好地描述广域网的技术特征,很简洁地给出了构成计算机网络的设备分类,对于计算机网络理论体系的形成产生过很好的作用,但是它不适用于描述结构复杂的互联网络。

OSI参考模型的7层结构的抽象方法实际上也是以广域网结构为对象,目的是希望统一当时各个大型计算机制造商在计算机网络体系结构模型出现时的混乱现象。图1-35给出了广域网结构与OSI参考模型的关系示意图。

但是,随着互联网的广泛应用,OSI参考模型的7层结构的抽象方法已经不能够满足复杂大系统的需要。我们可以用现实中的一个例子说明这个问题。如果我们分别是位于天津的A大学与位于成都的B大学的研究生,正在通过中国CERNET共同完成一项智能医疗

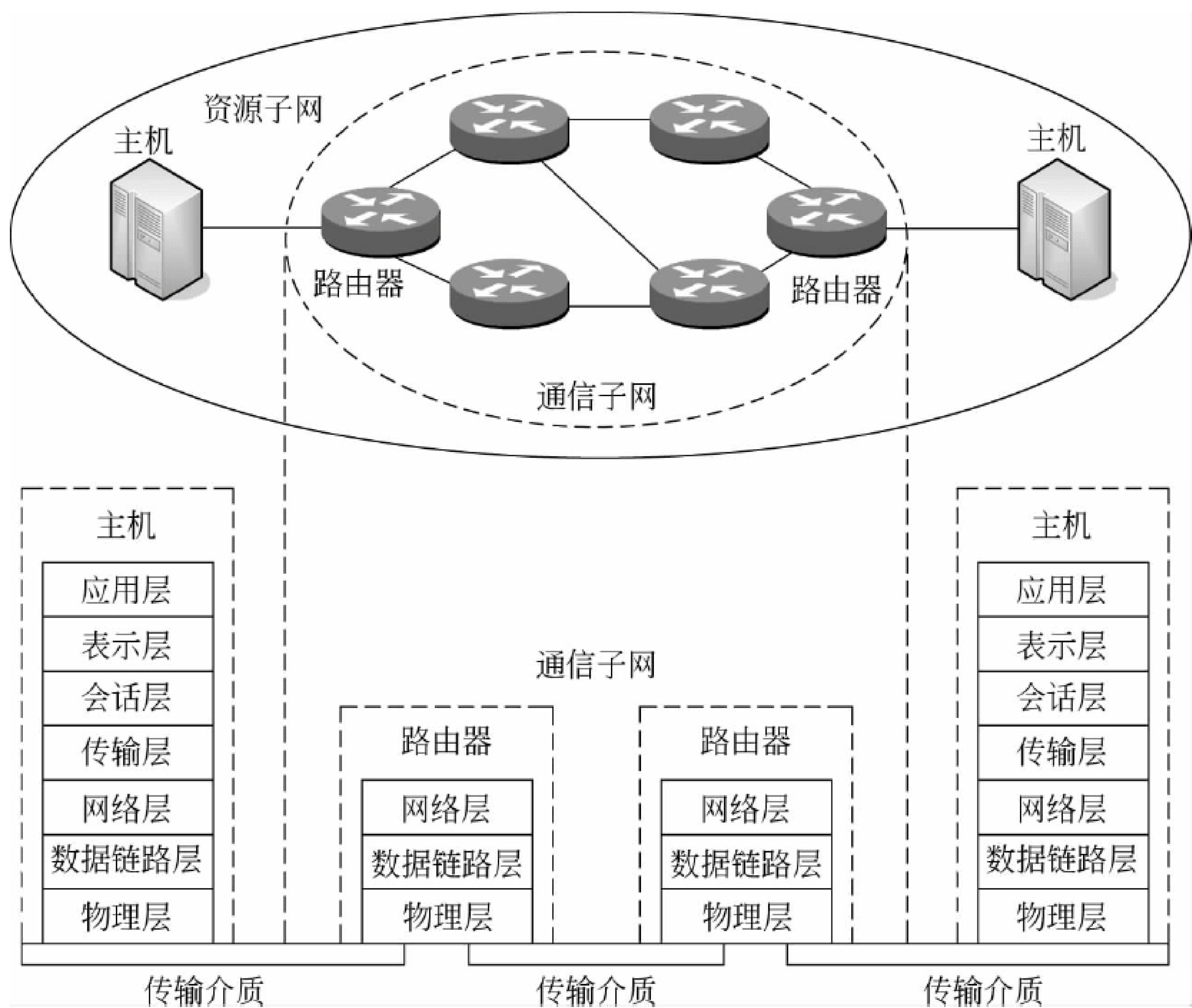


图 1-35　广域网结构与 OSI 参考模型

的合作研究，那么我们实际上是在如图 1-36 所示的计算机网络环境中工作。

从图 1-36 可以看出，工作在天津的 A 大学 A 实验室的研究生 A 与工作在成都 B 大学 B 实验室的研究生 B，可以通过计算机网络共享智能医疗合作研究项目中，分布在天津与成都的无线人体传感器网络 $WBSN_A$ 节点 A_m 与 $WBSN_B$ 节点 B_n 的数据。他们在相互交流实验数据，讨论用不同数据挖掘算法对数据分析结果时，好像就在一个实验室里"面对面交谈"一样。他们无须知道支撑他们交换数据的网络拓扑是什么样的，两台计算机之间交换数据使用的协议是什么样的，两台计算机之间数据分组是通过什么样的路由传递的，也不需要知道两台计算机之间进程通信交换数据的过程是什么样的。他们只希望知道：计算机之间交换的数据是正确的，协同工作中的数据"会话"与"交互"过程是流畅的，网络环境对于用户是"透明"的。这对于所有工作在网络环境中的科学研究人员来说，应该被看做是"理所当然"的事，而对于从事网络技术研究的技术人员，这正是他们希望实现的运行效果。但是，实际的网络工作过程远比人们想象的复杂得多。

天津的 A 大学研究生 A 使用的计算机，可以通过有线的 Ethernet 局域网或无线局域网 Wi-Fi 的方式接入到实验室 A 的局域网中；用于智能医疗项目研究的节点 A_m 的多种可穿戴医疗设备与传感器通过无线人体区域网 $WBSN_A$ 互联起来，再通过网关节点接入到实验室 A 的局域网中；实验室 A 的局域网通过路由器接入到大学 A 的校园网中。大学 A 的校园网通过路由器与交换机将学校的各个实验室、教室、图书馆、学生宿舍、办公室的成百上千个局域网互联起来，构成校园网，然后再通过校园网主干路由器连接到天津 CERNET。天津 CERNET 将天津几千所大学、中学、小学通过路由器互连起来，构成覆盖天津地区教育、科研单位的城域 CERNET。天津城域 CERNET 通过主干路由器接入国家 CERNET

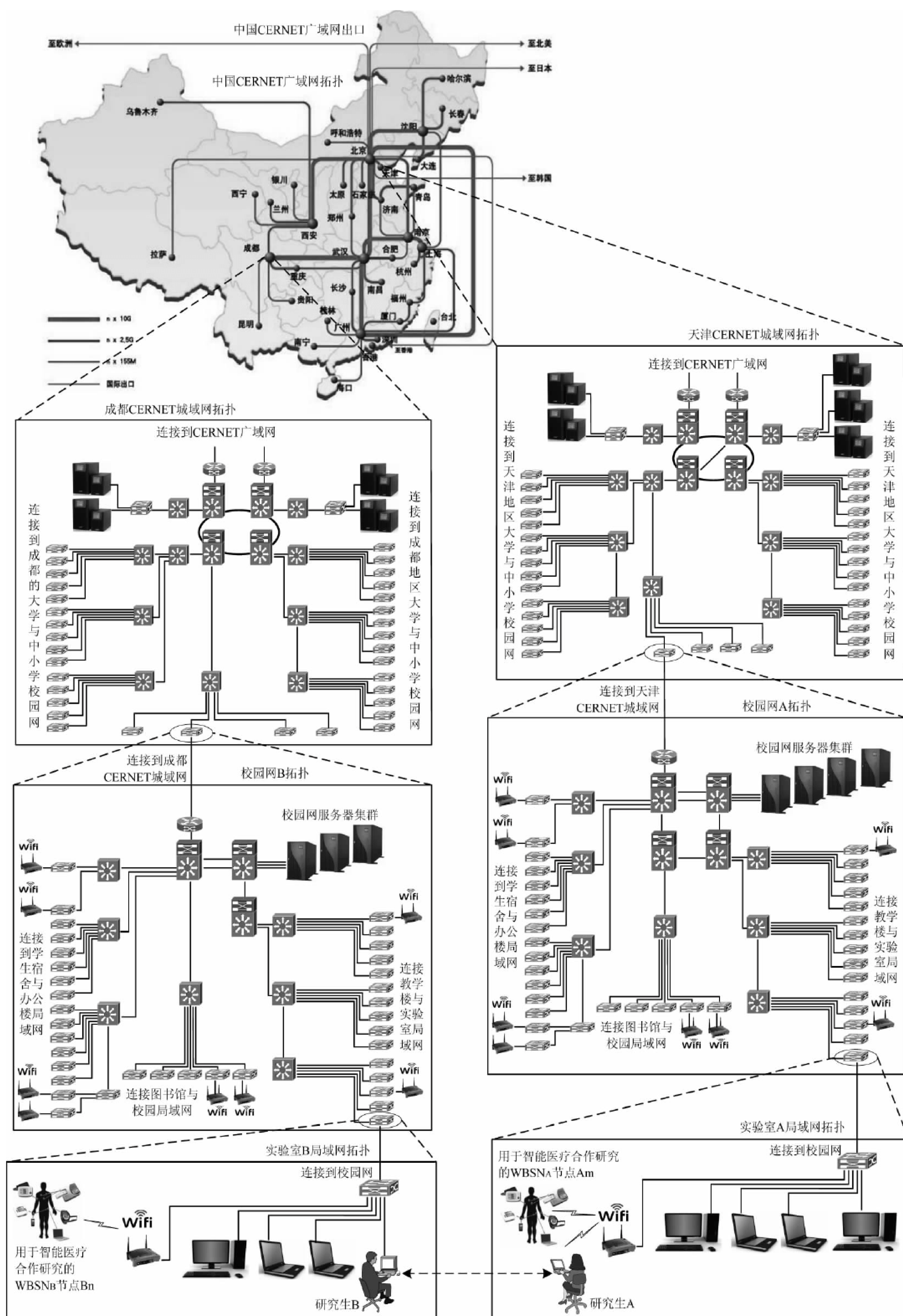


图 1-36　支持大学合作研究的我国 CERNET 网络结构示意图

主干网。国家CERNET主干网是一个覆盖全国的广域网。同样,成都B大学实验室B的计算机也会按照这样的连接方式,接入到国家CERNET主干网中,构成一个按层次结构连接、覆盖全国的大型教育科研网系统。

这里是以支持国内两所大学实验室合作研究的CERNET为例,来说明目前支撑我国大学教学、科研工作的网络环境。实际上我国有多个广域网,比如说中国电信的广域网、联通的广域网、铁通的广域网。我国多个广域网之间实现了互联互通,再通过我国国际互联网出口,与国际Internet主干网连接,使得无论是接入CERNET的科研教学用户,或者是中国电信、联通广域网的企业或办公室用户;无论是通过计算机网络方式,通过移动通信网3G/4G、电话交换网PSDN,或者是通过有线电视CATV网络接入的普通家庭用户;无论是通过固定的PC接入的用户,或者是通过iPad、PDA、智能手机、可穿戴计算设备接入的用户;无论是人或者是物(如传感器、RFID、智能机器人、嵌入式测控设备,以及车载网中的汽车),他们都能够接入到Internet,实现数据共享与协同工作。

研究复杂的网络系统时,可以采用"化繁为简"的抽象方法。图1-37给出了以我国CERNET为代表的复杂网络系统的层次结构示意图。

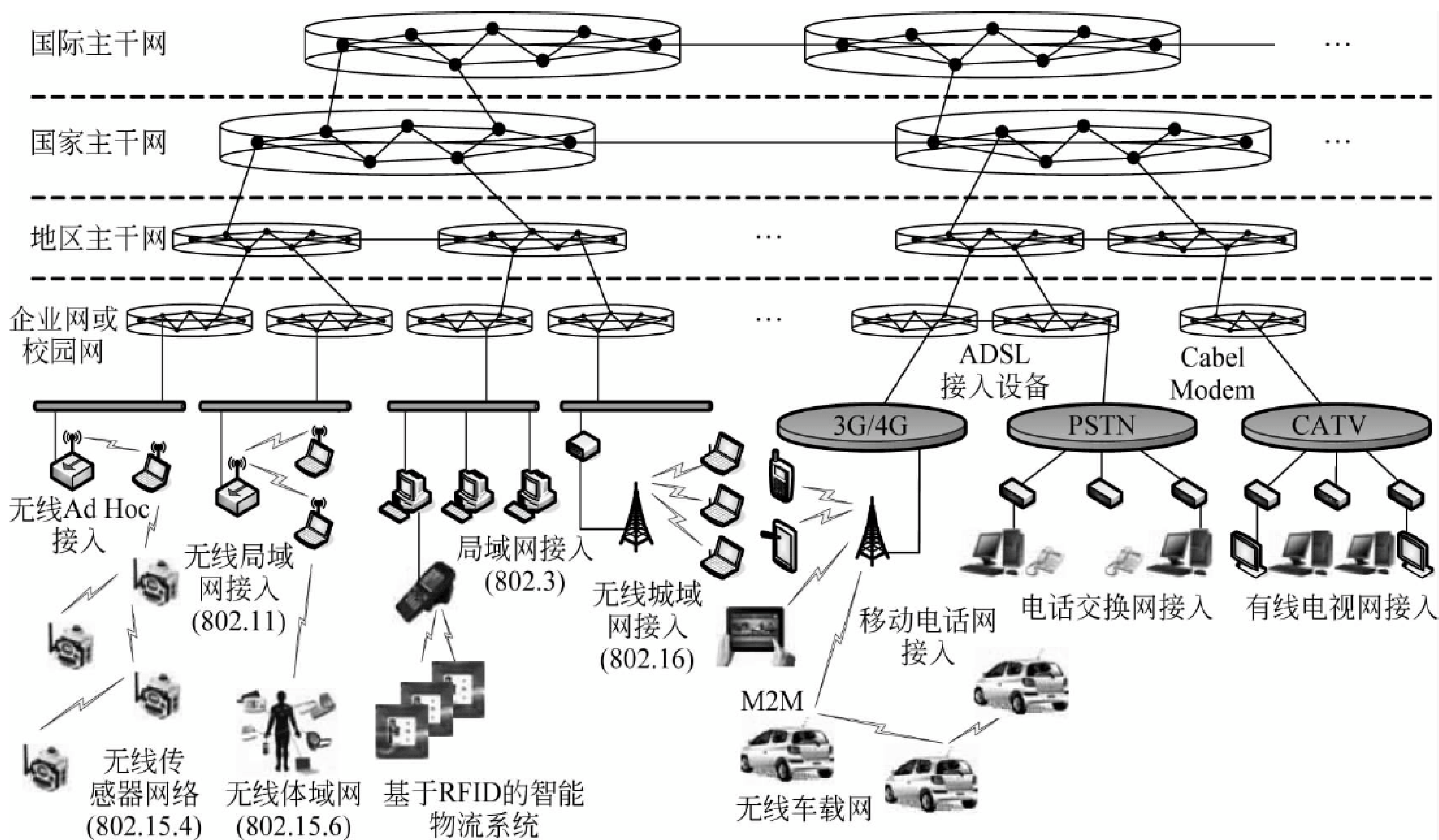


图1-37 层次化的网络结构模型示意图

图1-37是以我国CERNET为主要研究对象,针对大学实验室实际面对的大型计算机网络系统所做的第一次抽象。但是如果将协作研究的范围扩大到国内大学通过Internet与美国大学合作研究,那么我们面对的网络环境远比前面描述的结构要复杂得多。我们很难再画出一张类似于图1-35的网络结构。但是,自顶向下的计算机网络分析和设计方法给我们提供了一个很好的思路。

自顶向下的计算机网络分析和设计方法是将一个大型的网络应用系统,分解为边缘部分(端系统)与核心交换部分(传输网)两大部分。构成边缘部分的端系统由接入网络的计算机、智能终端设备组成,主要功能是通过分布式进程通信完成网络服务功能。核心交换部分

的传输网主要是由路由器与连接路由器的传输介质组成,主要功能是为应用软件的进程通信提供数据传输服务。图 1-38 给出了基于自顶向下的分析和设计方法对大型网际网结构所做的第二次抽象。

自顶向下的网络结构抽象描述方法可以很好地描述复杂的 Internet 结构,其中传输网包括:计算机网络中的广域网 WAN、城域网 MAN、局域网 LAN、个人区域网 PAN 与人体区域网 BAN,以及电信的移动通信网 3G/4G、电话交换网 PSTN、电视传输网 CATV。因此,Internet 传输网是由多种异构的网络互联起来的网际网。

基于自顶向下的分析和设计方法对 Internet 所做的抽象模型具有以下几个优点。

1. 使得复杂大系统的描述变得很简洁

随着广域网、城域网、局域网、个人区域网、人体区域网以及各种接入技术的发展,以及网络规模与应用类型的快速增长,我们实际面对的 Internet 结构越来越复杂,这就促使我们要研究适应 Internet 结构与网络应用系统的系统设计的描述方法。实践证明,“自顶向下”的分析与设计思路对于解决互联网应用系统设计与应用软件开发是非常有效的。依据“自顶向下”的分析与设计思路,人们可以将结构复杂、规模很大的 Internet 划分为“端系统”与“传输网”两大组成部分,并提出了“网络应用程序体系结构(Network Application Architecture)”的概念。端系统必须具备执行从应用层到物理层协议的能力,传输网中的路由器需要具备执行网络层、数据链路层与物理层协议的功能。这种抽象思维的方法抓住了事物最主要的特征,使复杂网络系统的描述变得简洁。

2. 使得网络系统的设计、实现与管理的界限变得很清晰

自顶向下的网络结构抽象描述方法对于网络系统的设计、实现与管理是十分有利的。网络应用系统设计人员的任务是:按照网络应用程序体系结构的思想,设计网络应用系统功能与结构,完成网络应用软件编程;利用传输网提供的数据传输服务,解决好互连的计算机之间分布式进程通信问题,实现预定的网络服务功能。网络运维与管理技术人员的任务是:运行与管理传输网中的路由器、通信线路,为连接在网络中的计算机之间的可靠数据传输提供技术支持。

3. 使得网络应用系统的设计、实现方法与步骤变得很清晰

按照“网络应用程序体系结构”的分析与设计方法,计算机网络与软件工程师在设计一个大型网络应用系统时,可以按照以下步骤开展工作。

第一步,根据应用需求,规划应用层功能,设计网络应用软件工作模式,选择应用层协议;根据应用层协议的要求,选择传输层是采用面向连接的 TCP,还是选择面向无连接的 UDP。

第二步,根据网络应用对数据传输的具体要求指标,选择适当的传输网的网络技术类型、结构和服务质量 QoS 指标,进而选择能够提供服务要求的传输网服务提供商。

第三步,依据应用层协议,开发网络应用软件。在完成网络应用软件编程之后,在实际的传输网环境中调试网络应用软件。网络应用系统测试通过后进入使用阶段。

第四步,网络应用系统在运行过程中,网络应用软件的维护、升级由计算机工程师负责;数据传输中存在的问题,由传输网服务提供商的通信工程师解决。

因此,按照“网络应用程序体系结构”的分析与设计方法,计算机工程师在设计一种新的网络应用时,他只需要考虑如何充分地利用核心交换部分的传输网所能够提供的服务,不涉

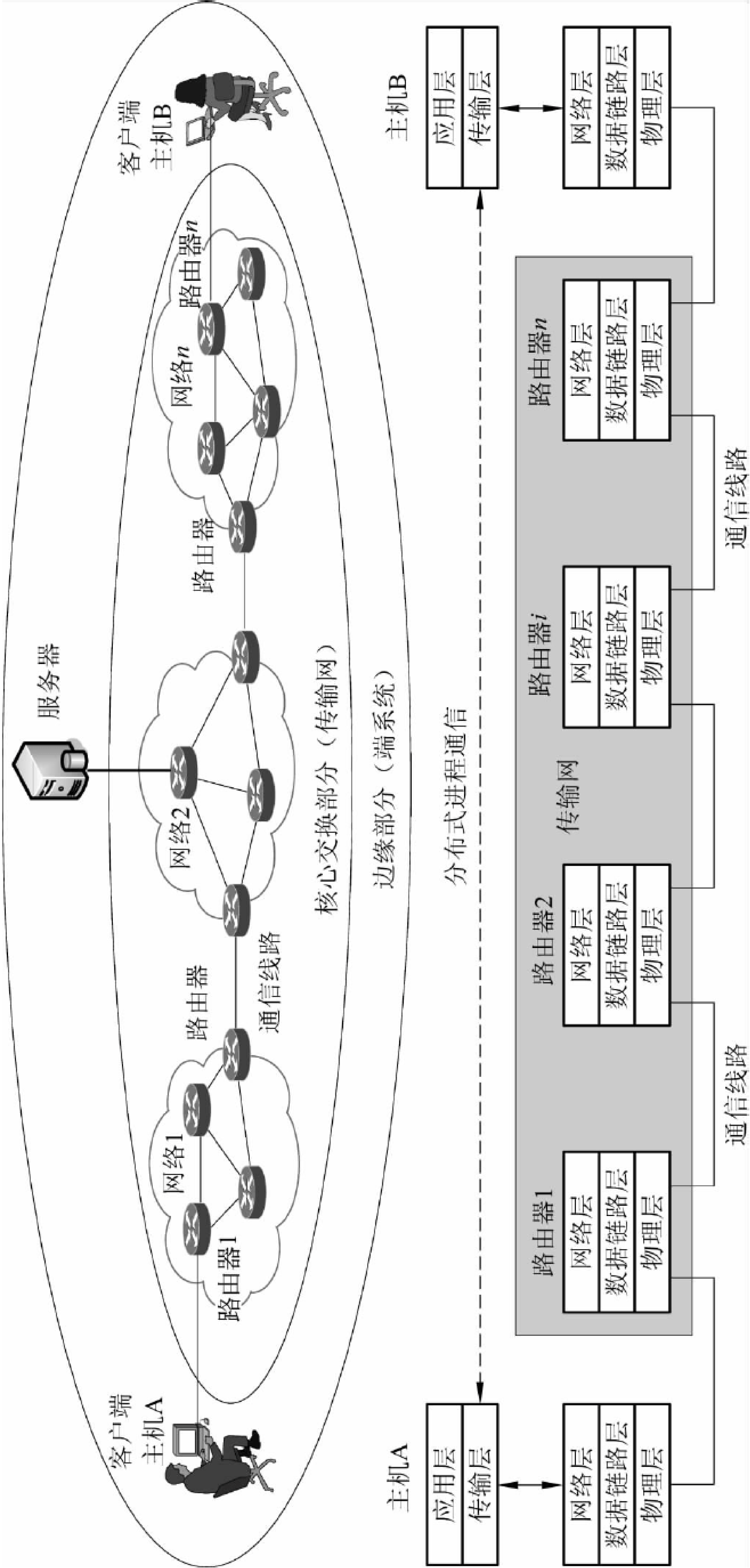


图 1-38 Internet 端系统与传输网的结构模型

及传输网中路由器、交换机等低层设备或通信协议软件的编程问题。这种分工明晰与密切协作的工作模式,保证了互联网、移动互联网与物联网各种网络应用系统可以快速地设计、开发与稳定地运行。

4. 使得互联网产业链的结构与分工变得清晰

一个成功的设计思想同时会使产业链的结构与分工明晰。在开发实际的网络应用系统中,除了有特殊需要之外,几乎没有任何一个人、任何一个单位、任何一个网络运营商、网络系统集成公司或软件公司,能够独立完成一个跨地区、跨国的大型网络应用系统,能够承担从规划、设计、软件开发到传输网的组建、运行管理的全过程的任务。跨地区的传输网一般都是由电信运营商、互联网服务提供商 ISP 运营。网络应用系统开发者在涉及使用广域网、城域网时,一般都是要租用电信运营商或 ISP 的通信线路与网络服务。传输网的日常运营、维护任务由电信运营商或 ISP 承担。

按照专业分工的思路,任何一个大型网络应用系统的设计者都是将复杂的问题"化整为零,分而治之",使设计、实施与运行、管理能够做到层次分明、功能清晰,使整个网络应用系统的设计与实现可以有条不紊地完成,网络系统可以长期、稳定地运行。在没有实际应用需求与财力、技术保障的前提下,任何一个想包揽大型网络应用系统开发、运行全过程的设想都是不现实的与不科学的,也是无法实现的。

问题 1-12:如何理解与应用 OSI 参考模型?

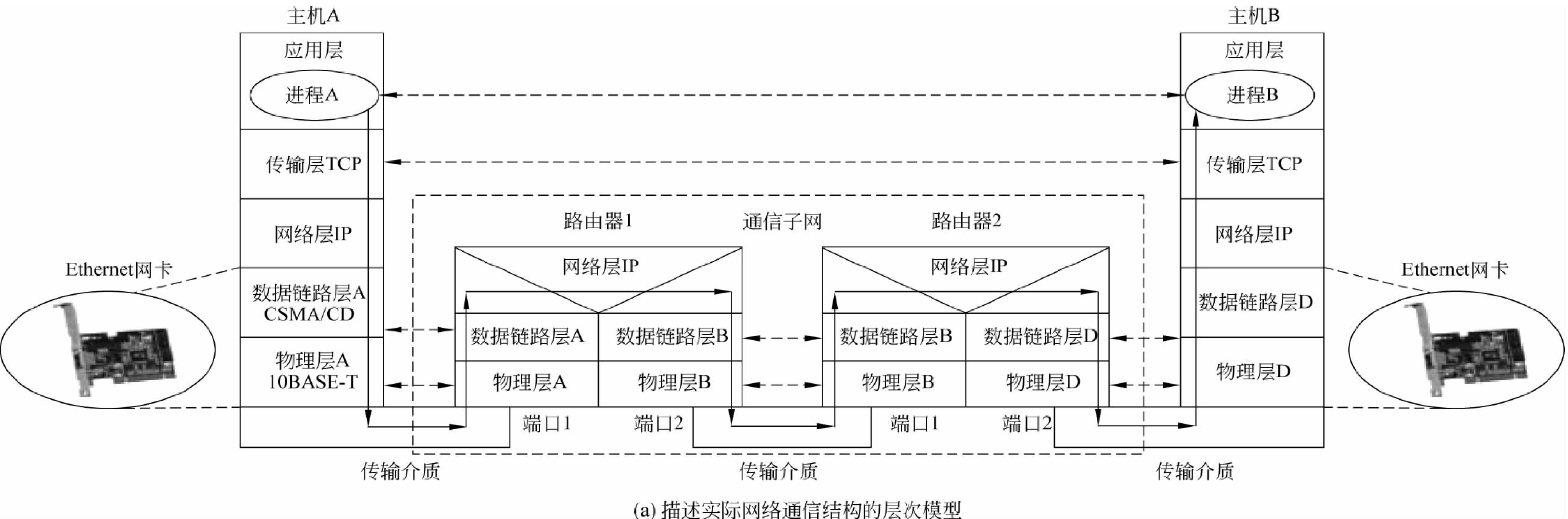
一般教材在描述 OSI 参考模型时都是采用如图 1-35 所示的结构画法。这是一个简化的模型,好处是可以使参考模型画得很简单。问题是对于初学者来说,在深入讨论网络工作原理时,常常会引起一些误解。如果将 OSI 参考模型和计算机网络中两台主机之间进程通信的数据传输途径联系起来,就可以画出如图 1-39 所示的 OSI 参考模型来。

理解 OSI 参考模型与主机进程通信的数据通信关系时,需要注意以下几个问题。

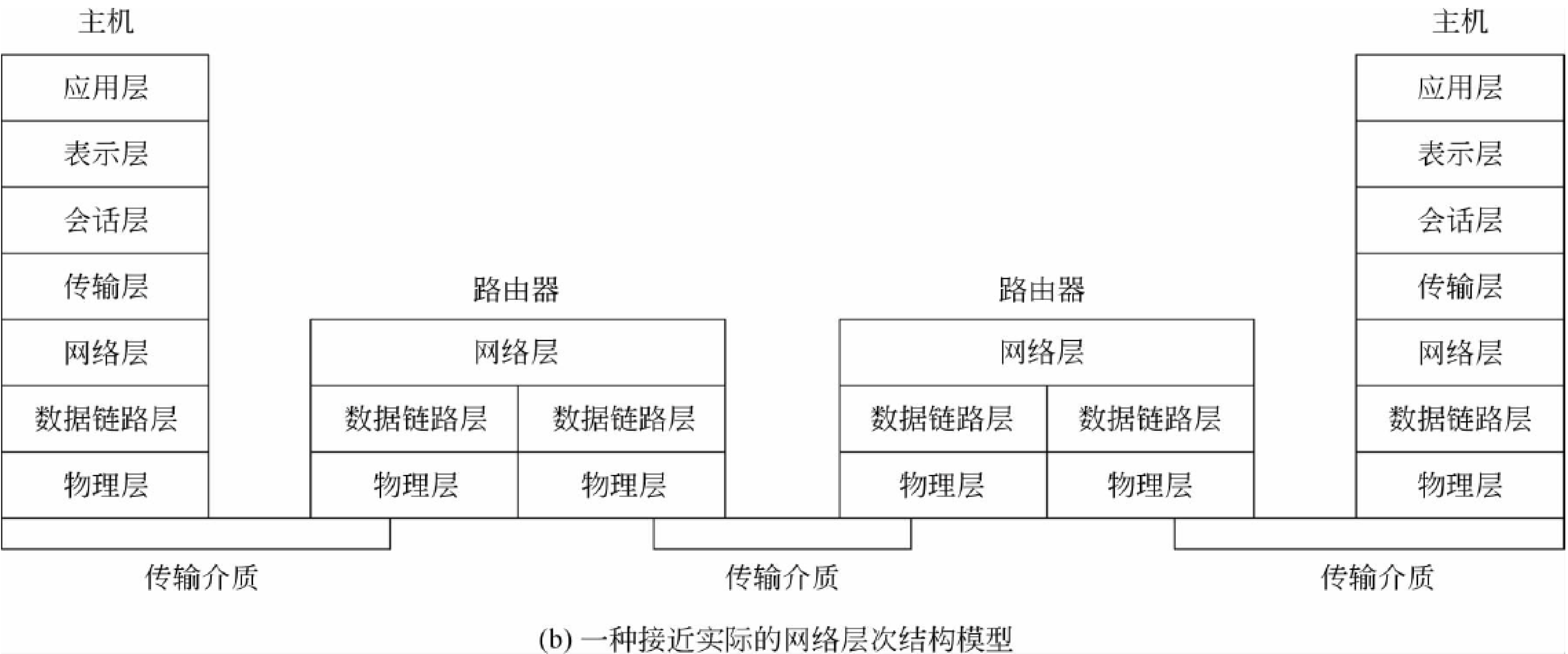
(1) 图 1-39(a)给出了根据典型的网络结构抽象出来的网络层次结构模型。我们可以用主机 A 的进程 A 与主机 B 的进程 B 的数据交互过程,来描述一种接近实际网络结构与通信过程的层次结构模型。

① 假设主机 A 与 B 采用了 TCP/IP 体系,那么主机 A 与 B 的传输层根据应用层协议的要求,选择 TCP 或 UDP。例子中的主机传输层采用的是 TCP。那么,无论是主机的网络层,或通信子网中路由器的网络层都要采用 IP。

② 尽管主机与通信子网中的路由器网络层都要采用 IP,但是它们的数据链路层与物理层可以灵活地选择不同的协议,只要相邻结点的对等层保持一致,就能够实现正常的通信。例如,主机 A 的常用 Ethernet 网卡接入到局域网,连接到路由器 1,那么主机 A 的数据链路层一定是采用了 IEEE 802.3 的数据链路层协议(CSMA/CD 协议)。而物理层有很多种 Ethernet 协议可供选择,主机 A 选用了传统 Ethernet 的 10BASE-T 协议。那么与主机 A 相邻的路由器 1 的端口 1 一端,也一定要选择与主机 A 系统的数据链路层(CSMA/CD 协议)与物理层(10BASE-T)协议。这样,主机 A 进程 A 的数据就可以通过传输层的 TCP 封装成 TCP 报文、IP 分组,再通过数据链路层封装成 802.3 协议帧。构成 802.3 协议帧的二进制比特流将按照 10BASE-T 协议的要求,变换成 Manchester 编码的电信号,通过网卡的 RJ-45 接口与双绞线,以 10Mbps 的速率传送到路由器 1 的端口 1。

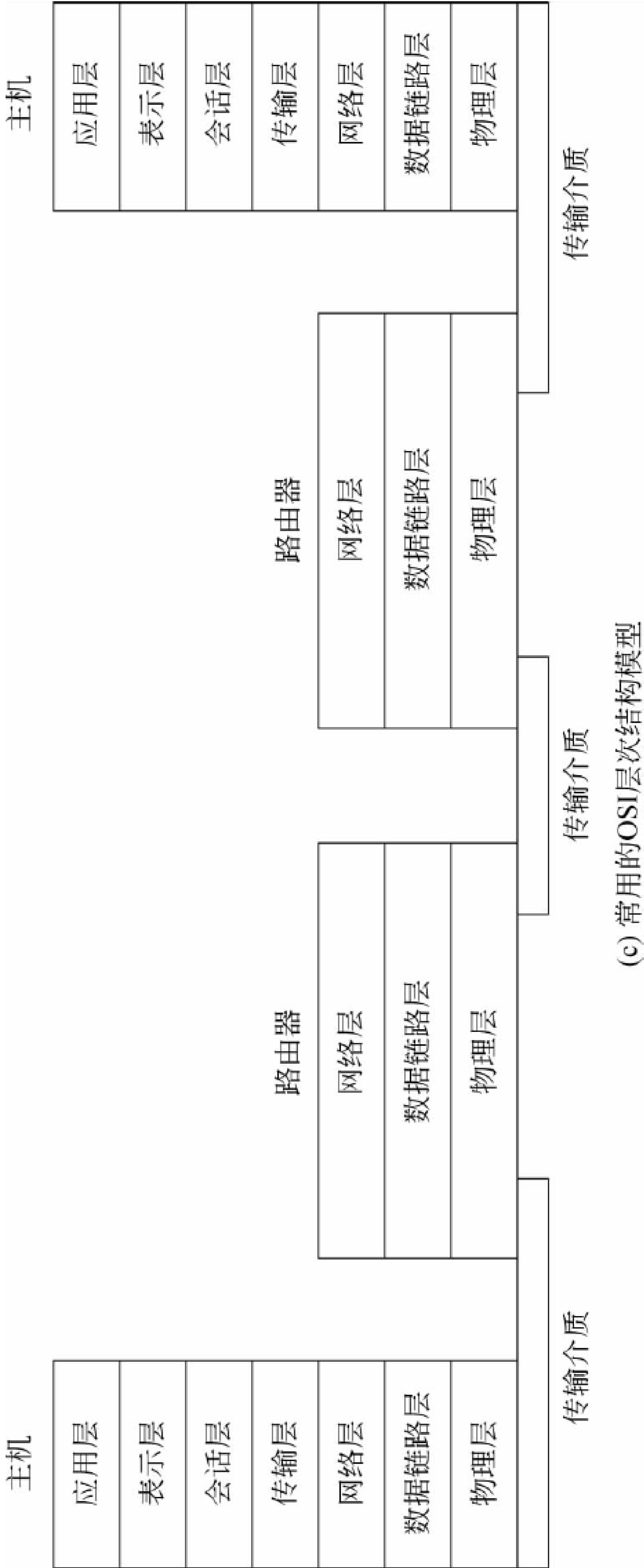


(a) 描述实际网络通信结构的层次模型



(b) 一种接近实际的网络层次结构模型

图 1-39 OSI 参考模型与主机进程通信的数据通信关系示意图



(c) 常用的OSI层次结构模型

图 1-39 （续）

路由器1的端口1与主机A连接，那么要保持与主机A正常的通信，路由器1端口1的数据链路层也必须采用与主机A系统相同的802.3协议，物理层采用10BASE-T协议。这样路由器A的端口1的物理层就可以正确地识别通过网卡的双绞线与RJ-45接口传送来的Manchester编码信号，传送到数据链路层，根据数据链路层的802.3协议对接收帧进行拆帧，取出IP分组传送到路由器的网络层。路由器的网络层启动路由选择算法，根据IP分组的目的地址找出下一跳路由器。

③ 如果下一跳路由器是路由器2，并且路由器1的端口2与路由器2的端口1连接，那么路由器1端口2的物理层、数据链路层就要和路由器2的端口1的物理层、数据链路层协议保持一致，但是不一定要与路由器1端口1的物理层、数据链路层协议一致。

路由器1端口2与路由器2的端口1的物理层可以选择光纤作为传输介质与相应的物理层光纤接口，数据链路层可以选择点-点的PPP。当然也可以选择无线信道作为传输介质，数据链路层选择无线网络的IEEE 802.11或802.16协议等。

为了简化通信子网的结构，层次结构模型的讨论中一般只选用了两个路由器。实际的计算机网络中通信子网有多个呈不同拓扑构型的多个路由器。它们的工作原理是一致的。

④ 主机A向主机B传送的进程A的数据通过通信子网之后，直到主机B的网络层才解析出TCP报文，传送到传输层；TCP软件继续解析出主机A发送来的应用层数据，传送给主机B的应用层进程B。直到这个时候，主机A的进程A与主机B的进程B之间的一次数据交换才结束。

这样的一个在计算机网络环境中分布式进程通信的过程可以用图1-39(a)的网络层次结构模型来表示。从作者多年的教学与科研实践中可以体会到，这种抽象的思考计算机网络体系结构的方法十分有用，它可以保证我们在网络软件编程中控制软件模块的规模与接口，可以用来解释各种新的网络应用系统的结构，可以解释移动互联网与物联网中很多大型应用系统的结构。

(2) 但是从计算机网络教学过程中，如果学生没有对计算机网络的工作原理与协议交互过程有一个深入的理解时，直接用图1-39(a)的结构去教学有一些困难，所以在教材中一般使用如图1-39(c)或图1-39(b)所示的结构。但是，在课程结束之前，一般要用1-39(a)的分析方法，解析几个实际的应用系统。这一点在物联网应用的教学中可以显示出很好的教学效果，对于提升学生应用网络技术应用能力至关重要。

问题1-13：如何认识预测互联网发展的新摩尔定律？

随着信息技术与互联网的发展，人们提出了10个预测性的定律，其中主要的4个预测性定律是：摩尔定律、吉尔德定律、麦特卡尔夫定律与新摩尔定律。

1. 第一个定律：摩尔定律

英特尔公司(Intel)创始人之一的戈登·摩尔(Gordon E. Moore)在1965年应邀为《电子学》杂志35周年专刊写了一篇题为“让集成电路填满更多的元件”的文章，对未来10年间半导体元件工业的发展趋势做出预言。他对收集的数据进行分析之后，发现了一个集成电路芯片集成度与时间关系的变化规律。经过1975年修正后表述为：“每过18个月，集成电路的性能将提高一倍，而其价格将降低一半”，也有人表述为“每过18个月，微处理机的处理速度将提高一倍”。这就是人们在描述信息技术，尤其是研究集成电路与计算机硬件技术发展趋势时，常常提到的“摩尔定律”。

计算机界从集成电路芯片集成度对计算机的计算能力影响的角度做出的推论是:“每过18个月,计算机的计算能力将提高一倍”。这就意味着每5年计算机运算速度会快10倍,每10年会快100倍。同等价位的微处理器越来越快,同等速度的微处理器越来越便宜。这个规律也适用于描述存储器的发展趋势。

2. 第二个定律:吉尔德定律

1995年,乔治·吉尔德曾预测:在未来25年,主干网的带宽将每6个月增加一倍。这就是吉尔德定律(Gilder's Law)。乔治·吉尔德认为,正如20世纪70年代昂贵的晶体管,在现如今变得如此便宜一样,如今还是稀缺资源的主干网带宽。如果有一天变得足够充裕的带宽,那时人们上网的费用将大幅度下降。

吉尔德定律所描述的主干网增长速度比CPU增长速度要快。只要将廉价的网络带宽资源充分利用起来,就会给人们带来巨额的回报。未来的成功人士将是那些更善于利用带宽资源的人。这个定律已被很多基于互联网的应用所证实。

3. 第三个定律:麦特卡尔夫定律

大约在1980年,Ethernet的发明人鲍勃·麦特卡尔夫(Bob Metcalfe)指出:网络的价值与网络用户数量的平方成正比。从目前互联网应用的发展情况来看,这个定律已经被Web、Facebook、Google、Blog与移动互联网所印证。

2013年1月17日,瑞典互联网市场研究机构Pingdom发布的全球互联网用户数为24亿,其中:亚洲11亿、欧洲5.19亿、北美地区2.74亿、拉美/加勒比海地区2.55亿、非洲1.67亿、中东9000万、大洋洲2430万、中国5.65亿。

截止到2015年年底,我国的互联网网民规模已经达到6.88亿,普及率达到50.3%。手机网民规模达到6.2亿,占网民总数的90.1%。这些数据都进一步展示出:互联网、移动互联网与物联网对未来社会发展将会产生越来越大的影响。

4. 第四个定律:新摩尔定律

实际上,我们所讨论的“新摩尔定律”是另一个对全球互联网发展规律的预测的定律。联合国“1999世界电信论坛会议”副主席、加拿大北电网络公司总裁约翰·罗斯(John Roth)在论坛开幕演说中提出了著名的“光纤定律(Optical Law)”。光纤定律指出:互联网通信速率每9个月会增加一倍,成本降低一半。人们经常将对互联网通信速率与成本的预测叫作“新摩尔定律”。目前,互联网的广域主干网、地区汇聚网,甚至是家庭接入网基本上都是采用光纤专线连接路由器的结构,就进一步证实了“光纤定律”预测的结论已经被产业界所接受。

新摩尔定律与前三个定律一样,它们都不是数学、物理定律,而是对技术发展趋势、规律的一种预测性的定律。在最近的几十年来,计算机、计算机网络与互联网的发展证实了这些预见的正确性。这些定律对于指导计算机、互联网与信息技术的发展有着重要的指导意义,因此受到了产业界与学术界的重视。

我们可以援引麦肯锡全球研究所(McKinsey Global Institute)在2013年公布的一份关于潜在的12项颠覆性技术对全球经济发展影响的数据来印证以上的分析意见。图1-40给出了12项颠覆性技术中排在前6位的技术的相关数据。这6项技术直接与网络技术相关,或是在网络技术支撑之下发展的技术。报告指出:在到2025年的未来10年中,对世界经济发展贡献排在第一位的是“移动互联网”,它每年将可以创造3.7万~10.8万亿美元的经

济效益;排在第三位的是“物联网”,它每年将可以创造 2.7 万~6.2 万亿美元的经济效益。

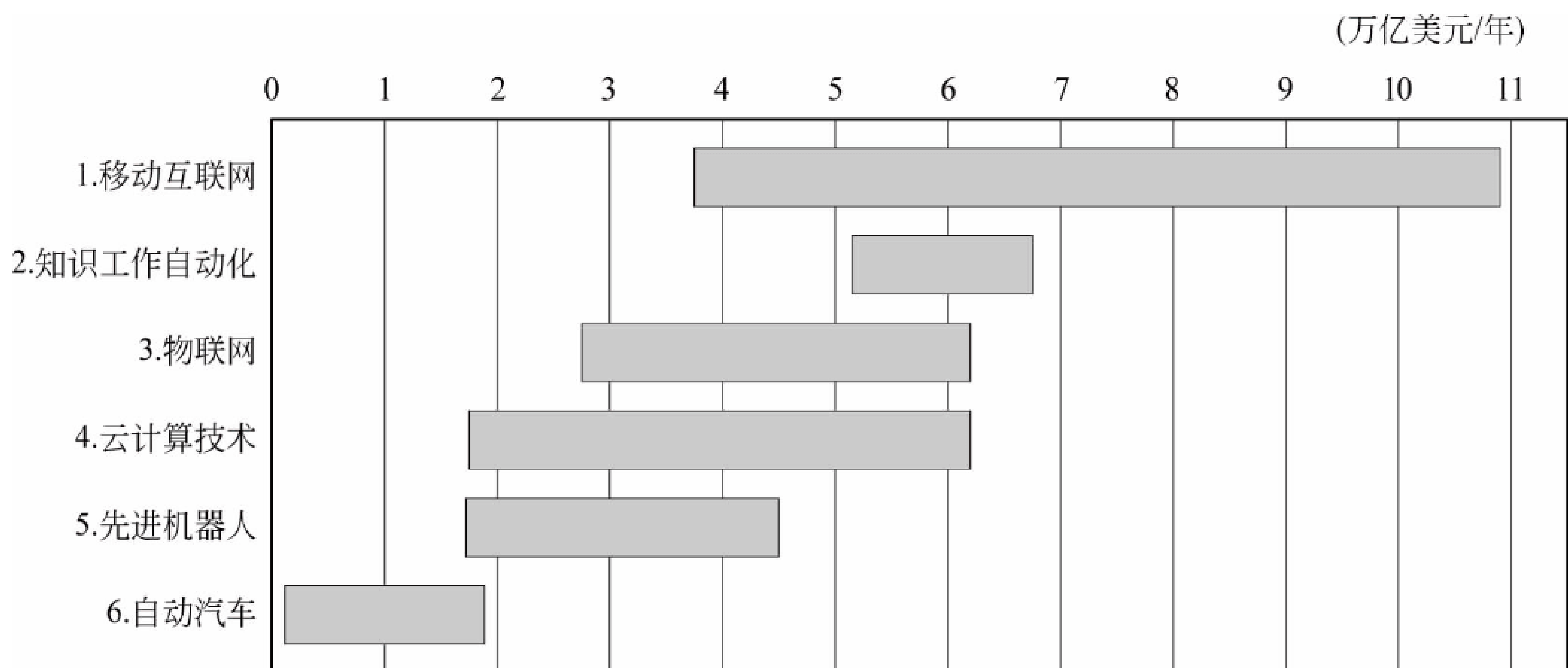


图 1-40　12 项颠覆性技术中排在前 6 位的相关数据

问题 1-14:如何认识互联网发展的成功经验?

讨论互联网发展成功经验的目的是希望对今后发展物联网提供经验与借鉴。互联网是人类历史上发展最快的一种信息技术。我们可以从一组数据来看出这个问题。

从开始商用到用户数达到 500 万,电话网用了 100 年,无线广播网用了 38 年,有线电视网用了 13 年,而互联网只用了 4 年。这一组数据说明:互联网技术是很成功的。

对于互联网发展成功的经验,早在 1996 年 6 月发表的 RFC1958(Architectural Principles of the Internet)中已经开展了研究。Tanenbaum 在《计算机网络(第 5 版)》关于互联网的讨论中总结出互联网设计的十大原则。

结合以上的讨论,反思互联网技术发展与演变的过程,可以将互联网成功的发展经验概括为:正确的设计思路,正确的技术路线,正确的运行模式。

1. 正确的指导思路

我们面对的互联网是一个异构、动态变化、复杂的网际网。假设用一台南开大学网络实验室的计算机访问 MIT Auto-ID 实验室一台主机的数据库时,谁也无法预测和控制这次通信的分组是通过哪条路径到达对方的。回顾互联网的发展过程,我们清楚地看出,面对复杂的互联网环境,网络先驱者采用了正确的设计原则,成功地解决了这个问题。他们所采用的设计原则是:

(1) 明确选择;

(2) 保持简单。

回顾 IP 协议研究与发展,总结初期 IP 协议讨论的内容、IP 协议整个改进和完善的过程,以及实际取得的效果,可以从中得出的启示是:IPv4 协议的设计是成功的,它为互联网的发展奠定了坚实的基础。它的成功表现在以下两个方面。

一是技术路线是正确性的。如果要求计算机科学家在 20 世纪 70 年代就能够预见计算机网络可能发展到今天这样大的规模,产生如此深刻的影响,那是不可能的。当时出现过各种设计方案,有的很简单,如只能提供“尽力而为”的 IP 服务。有的很复杂,他们预先要考虑各种可能,提出一系列应对策略与复杂的协议体系。例如,OSI 参考模型与协议文本堆起来

有一米高。当选择IP协议的技术人员已经开始设计路由器、开发软件时,选择OSI参考模型与复杂协议方案的技术人员还没有看明白协议文本,更谈不上进入硬件设计与软件编程了。互联网设计者采取了"明确选择""保持简单"的处理原则,促进了互联网的发展。

IP协议的设计者在第一个设计文档中只对IP分组结构做出了规定,对IP地址按照标准分类的方法给出了意见,提出了直接交付与间接交付、路由选择的概念。他们采用简单的方法去解决复杂问题,用"尽力而为"的服务去应对互联网络中可能存在的各种复杂问题。这样做才有利于技术的推广与应用,才在TCP/IP体系与OSI参考模型的竞争中赢得了时间与市场,吸引了大批资金的投入。

二是伴随着互联网规模的扩大和应用的深入,作为互联网核心协议之一的IPv4协议也一直处于一个不断补充、完善和提高的过程,但是IPv4版本的主要内容没有发生任何实质性的变化。实践证明,IPv4是健壮和易于实现的,并且具有很好的互操作性。它本身也经受住了互联网从小型的科研范围应用的互联网络,发展到今天这样的全球性大规模网际网的考验,这些都说明IP协议是成功的。

这种"化整为零,分而治之"的分层设计思想指导了整个分布式互联网的体系结构与协议体系的设计。例如,在域名与域名服务DNS体系的设计,自治系统与EGP/BGP路由协议的设计,网络管理与网络管理协议SNMP的设计中都能够充分地体现出"保持简单"的设计思想。

事实证明,简单的、无连接的、"尽力而为"的IP协议就是最恰当的选择。"尽力而为"的IP协议设计思想,它简化了网络协议的设计与实现,同时也提高了网络系统的可靠性。TCP/IP体系的成功显示了互联网设计者"明确选择""保持简单"设计原则的正确性。

2. 正确的技术路线

互联网技术路线的正确性表现在以下两方面。

(1) 选择好的设计,而不是完美的;

(2) 考虑性能与成本。

凡是早期参与计算机网络路由选择协议研究的网络技术人员都有一个亲身体会,那就是:复杂网络的路由选择算法研究的难度非常大。20年前,每年的IEEE网络年会都会有几十,甚至上百篇关于路由选择算法的论文,涉及的算法从简单路选算法,到复杂的自适应路由选择算法都会有。但是细细读后就会发现,每一篇都有道理,但是都不可能适应所有不同的情况。然而,在实际互联网工程实践中,只要引入了"自治系统"的概念和内部网关协议与外部网关协议,只用选择RIP、BGP或OSPF等几种协议,尽管每一种路由选择协议都有自己的适用范围与性能的限制,但是我们完全可以采用简化的方法,去构建目前广泛适用的、可扩展的互联网系统。显然,这种做法充分体现出"选择好的设计,而不是完美的"的技术路线。

"简洁、实用"是互联网选择技术的基本原则。例如,目前互联网主干网结构中基本上都是采用路由器加专线的互联方式,专线主要是使用光纤。这样做的好处是:既简化了传输网的结构,又有利于提高网络系统的可靠性与性能。同时,在端-端进程通信中采取高层协议解决流量控制、拥塞控制的方法,这样做可以减轻传输网的压力,合理地分配了保证分组传输可靠性处理的负荷,提高了网络系统整体的效率。

互联网的发展促使了计算机网络、电信网络与电视网络在技术与业务上的三网融合,带

动了信息通信产业的发展转型。在全社会“三网融合”艰难推进之时,计算机屏幕、手机屏幕与电视屏幕已经在人们的手掌中率先实现了“三屏融合”,反过来又促进了三网融合的发展。

3. 正确的运行模式

开放性、社会性与可扩展性是互联网运行模式的重要特点。互联网开放性首先表现在:互联网不属于任何公司与个人所有,而是由非营利的组织(如 ISOC、IETF、IRTF 等)、行业组织(如 W3C、Wi-Fi Alliance 等)参与协议标准制定和产业发展指导。应用驱动、开放合作的研发模式,是互联网得以超常规发展的重要基础。

互联网的社会性表现在:人与人处于虚拟的、不是直接见面的环境中交流,这种交流可以克服年龄、职业、地位、性别与性格上的差异,尽情释放人性中最自然的一面,使得互联网服务具有它特殊的魅力。互联网服务可以克服现实生活人与人之间时间、空间限制,使得世界变得很小,使得人们的生活更加丰富多彩,人与人、人与社会的沟通更加便捷。

互联网的可扩展性表现在:统一技术标准,集成一切可用的技术,鼓励通用的应用技术开发。按照“用户需求—技术研究—标准制定—产品研发—产业发展”的发展思路,这是互联网技术与产业遵循的发展规律。互联网产业链是由网络硬件制造业、网络软件业、网络运营业、基于互联网的现代信息服务业构成。网络硬件制造业包括网络设备制造业、电子信息产品制造业;网络软件业包括网络系统软件与应用软件的研发企业;网络运营业包括电信网络运营商、有线电视网络运营商,以及计算机网络运营商;基于互联网的现代信息服务企业包括互联网服务提供商(ISP)、互联网内容提供商(ICP)与互联网应用提供商(IAP)。

问题 1-15:术语辨析:Computer Network、internet、Internet 与 Intranet。

(1) Computer Network(计算机网络)表述的是互连的、独立计算机系统的集合。计算机网络有各种类型,如广域网、城域网、局域网或个人区域网。

(2) Internet 或 Internetworking(网络互联)表述的是一种互联多个计算机或计算机网络的技术。

(3) Internet 也称为互联网、因特网,它是一个专用名词,专指目前广泛应用、覆盖了全世界的、由很多网络互联而成的网际网。

(4) 随着 Internet 的广泛应用,一些大型企业、管理机构也采用了 Internet 的组网方法,采用 TCP/IP 与 Web 的系统设计方法,将分布在不同地理位置的部门局域网互联成企业内部的专用网络系统,供内部员工办公使用,不连接或不直接连接到 Internet,这种内部的专用网络系统叫作 Intranet。

问题 1-16:术语辨析:结点与节点、互连与互联。

1. 结点与节点

英文“node”中文可以译为“节点”或“结点”。在中文教材、译著或学术论文中,很少注意两者的区别,有的教材在所有涉及“node”术语的地方都用“节点”或都用“结点”表述。作者注意到在网上也有关于这个问题的讨论。一种意见是:两线相交,其相交点属于中间点的叫作“节点”。两线相交,其相交点属于终点的叫作“结点”。也有的意见是借鉴电路课或 C 语言的概念。另一种意见是不区分,都用“节点”或“结点”。

作者认为作为计算机网络的中文教材,还是应该注意用中文表述技术术语时,在内涵上的细微区别与准确性。在主教材中,读者可以注意一个细节。依照拓扑学的规定:凡是将

网络中的实体抽象成与其大小、形状无关的点，统称为“节点”。这种情况可能出现在图 1-41 描述的问题中。例如，在研究计算机网络拓扑时(图 1-41(a)所示)，在这种问题中无论它是总线结构、环状结构或者是网状结构，网络中的实体都被抽象成一个点，不需要讨论这些点在物理结构上的区别。这种情况还可能出现在如图 1-41(b)所示的网桥生成树算法，以及如图 1-41(c)所示的简单网管协议 SNMP 的对象命名树的讨论中。当然，可能还不止这些情况。但是，判断是否用术语“节点”描述的原则是：将网络中的实体抽象成一个点，在研究过程中不需要考虑被抽象的设备的类型、配置与功能。

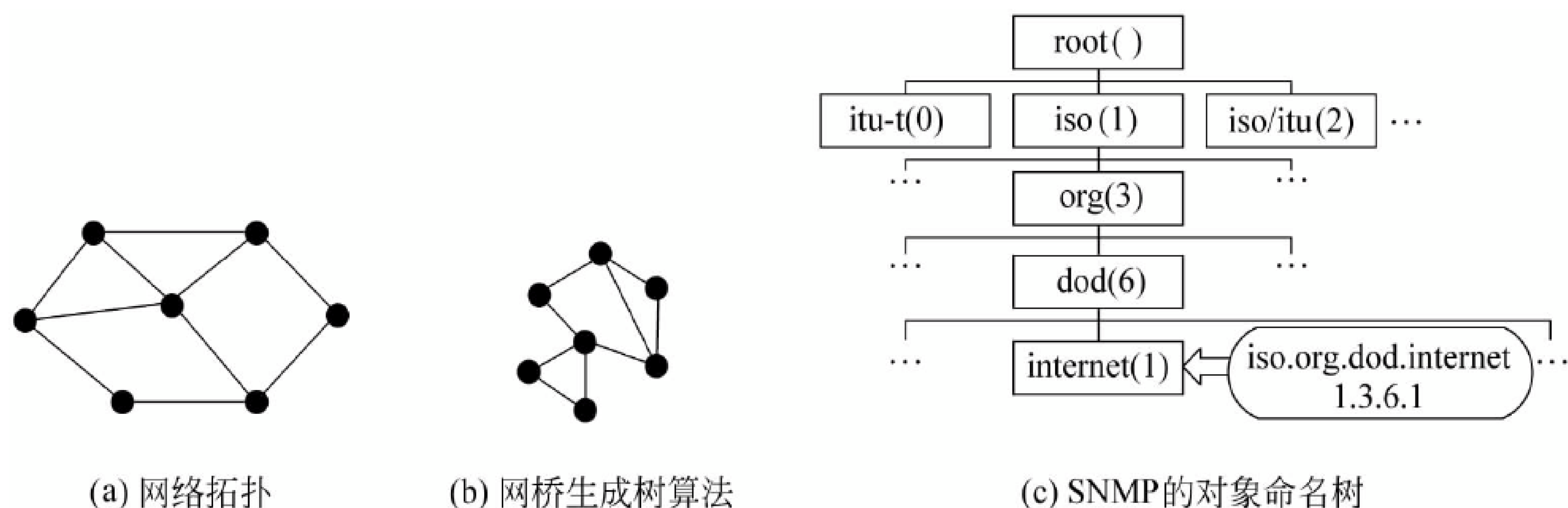


图 1-41　抽象为“节点”的几种情况

反之，如果在研究过程中需要将网络中的实体做一定的抽象，同时要考虑被抽象设备的类型、配置与功能。这种情况如图 1-42 所示。例如，在研究 Ethernet 工作原理时(图 1-42(a)所示)、研究路由协议(如图 1-42(b)所示)时的路由器、研究传输层与应用层网络应用与应用进程通信时的客户与服务器(如图 1-42(c)所示)，我们通常将这种情况下的计算机、路由器简称为“结点”。当然，在很多场景下也可以将网络中的计算机称为“主机(Host)”。

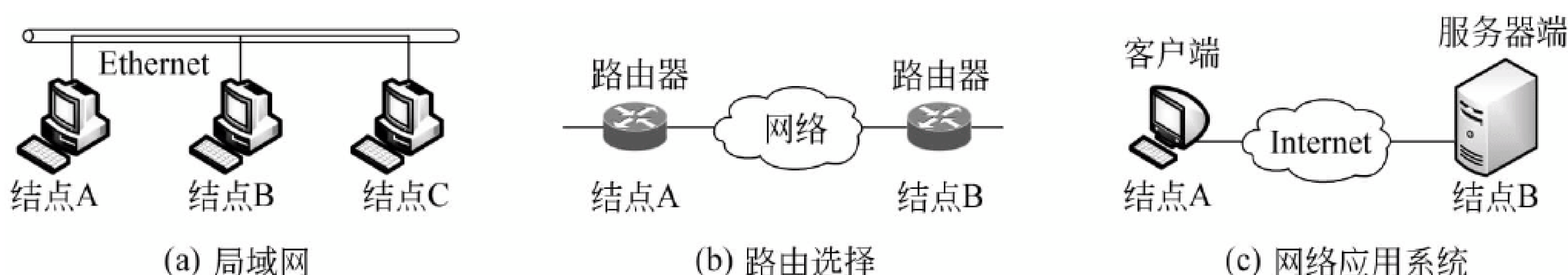


图 1-42　抽象为“结点”的几种情况

2. 互连与互联

互连与互联是网络教学中经常抽象的技术术语，两者是有区别的。

互连(Interconnecting)强调的是计算机与计算机、计算机与交换机、计算机与路由器之间的物理连接。例如，在局域网结构化布线技术中，我们重视的是设备之间的物理连接，相互之间可以传输比特流，而不强调它们能实现应用程序之间的进程通信。如果说计算机网络实现了网络中计算机之间的“互连、互通、互操作”，那么术语“互连”只强调“互连、互通”。而互联(internetworking)则更加强调的是计算机之间在互连、互通的基础上，能够实现互操作(interoperation)。因此，术语“互连”用于描述物理层的问题，而术语“互联”用于描述数据链路层及以上高层的问题。

问题 1-17：你能不能对网络课程讲授的技术做一个综述？

我们用一张网络课程需要讲授的主要技术之间关系的图来表述，可能会更简洁、明了。

图 1-43 给出了网络课程讲授的主要技术结构示意图。

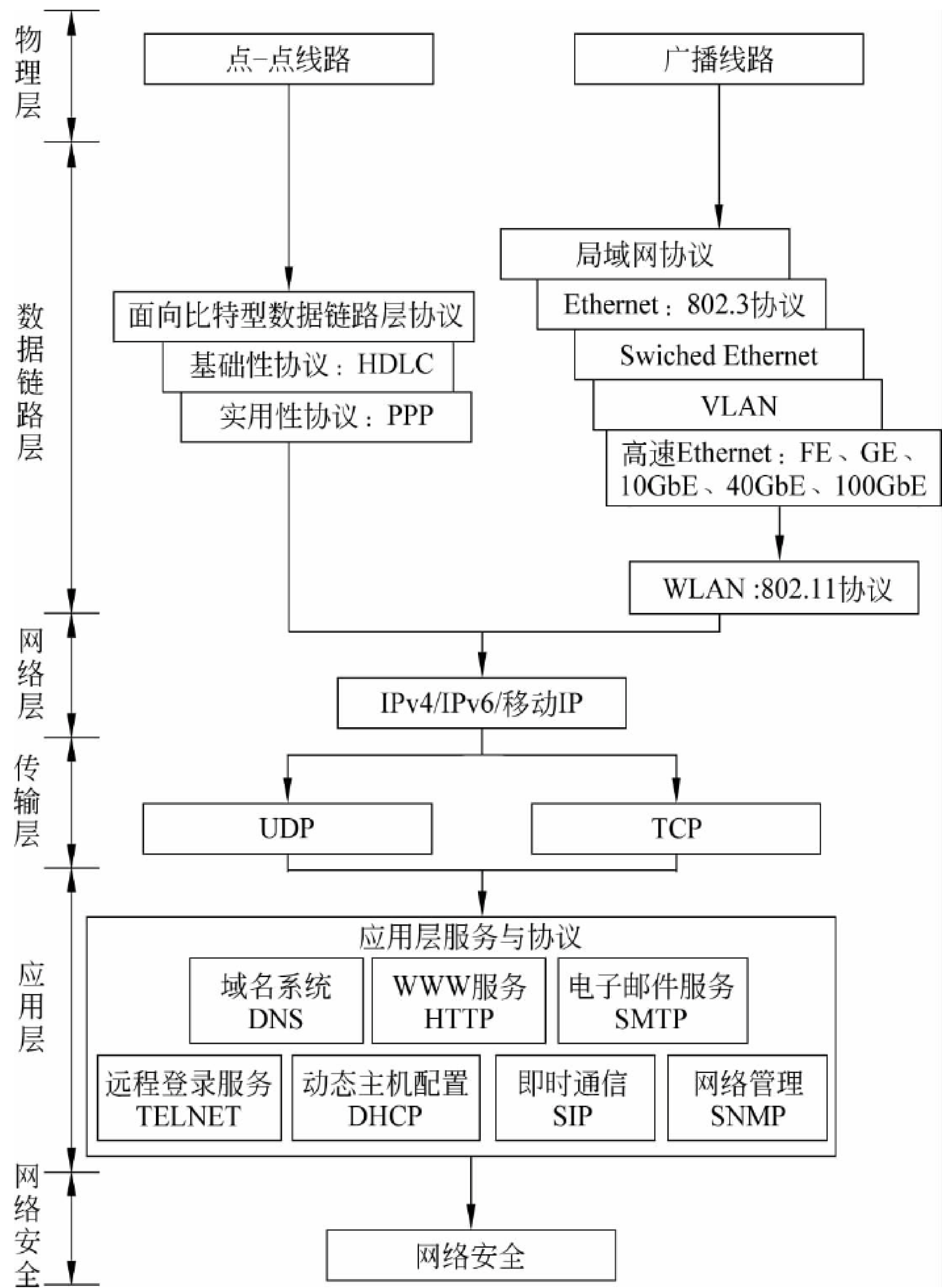


图 1-43　计算机网络课程讲授的主要技术结构示意图

第三部分　习题参考答案

5.（1）长度为 8B 的传输效率约为 7.5%。

（2）长度为 536B 的传输效率约为 85.9%。

6.（1）发送延时为 10ms；传播延时为 5ms。

（2）发送延时为 100ms；传播延时为 5ms。

7.（1）报文交换总延时约为 2255.3ms。

（2）分组交换总延时约为 290.5ms。

10. RFC791 文档

名　　称：INTERNET PROTOCOL

发表时间：September 1981

第 2 章

物理层

第一部分　学习目的、要求与知识点结构

1. 学习目的

数据通信技术是网络技术发展的基础。本章将对数据通信的基本概念、主要传输介质、数据编码技术、数据传输技术、多路复用技术与差错控制技术进行系统的分析。学习本章内容将对网络中最基本的数据通信技术、广域网中数据传输原理与实现方法的理解有很大的帮助,为以后的学习打下坚实的基础。在教学过程中需要注意的是,学习物理层知识对于理解网络基本工作原理是非常重要的,但是物理层的很多内容涉及的是通信技术的细节问题,因此在计算机网络课程的学习中属于广度优先的章节,学时安排上应该适度。

2. 学习要求

(1) 理解:物理层与物理层协议的基本概念。

(2) 理解:数据通信的基本概念。

(3) 掌握:传输介质类型及主要特性。

(4) 掌握:数据编码的类型和基本方法。

(5) 掌握:基带传输与频带传输的基本概念。

(6) 掌握:多路复用技术的分类与特点。

(7) 掌握:同步数字体系 SDH 的基本概念。

(8) 掌握:接入技术的基本概念。

3. 本章知识点的组织与结构

本章知识点的组织与结构如图 2-1 所示。

本章知识点的组织中,作者注意了以下几个问题。

(1) 广域网主要采用点-点通信线路,局域网与城域网一般采用广播信道。由于广域网、局域网和城域网所采用的通信线路类型不同,并且广域网研究最早,在广域网技术研究的基础上,人们才开始局域网技术的研究,城域网基本采用与局域网相同的技术,因此广域网与局域网、城域网之间在技术特点上存在着较大的差异。正是有这样一个技术发展过程,因此在计算机网络的物理层和数据链路层协议上出现了两个分支,一类是基于点-点通信线路的广域网物理层和数据链路层协议,另一类是基于广播信道局域网与城域网的物理层和数据链路层协议。

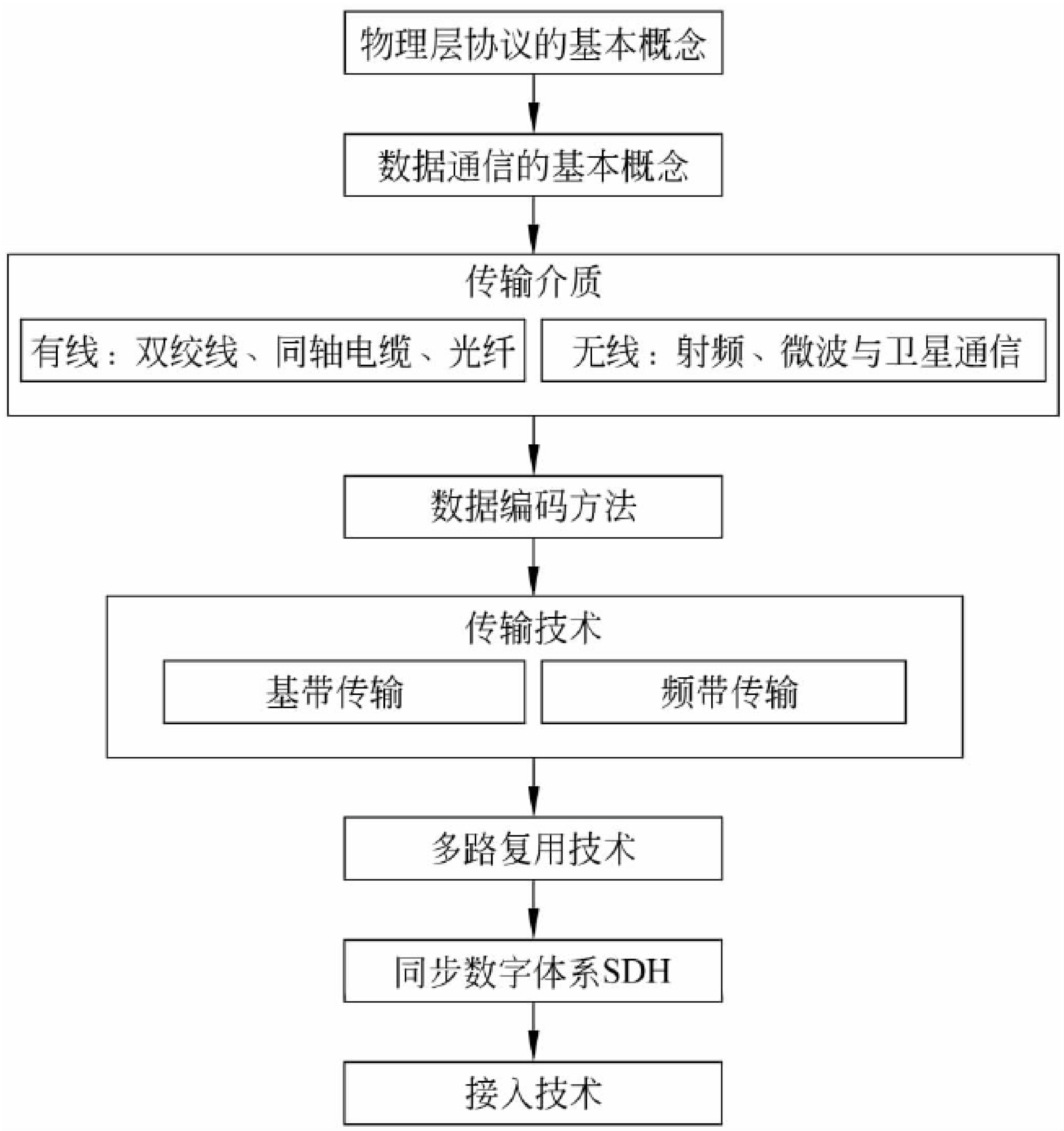


图 2-1　知识点的组织与结构

（2）基于点-点通信线路的广域网物理层和数据链路层技术与协议的研究开展得比较早，形成了自己的体系、协议与标准。基于广播信道的局域网（Ethernet、Wi-Fi）的物理层和数据链路层协议研究相对比较晚一些。人们根据局域网的技术特点，对 OSI 参考模型中的数据链路层做了相应的修改，提出了介质访问控制子层与逻辑链路控制子层的概念，研究了基于广播信道通信特点的体系、协议与标准。在这些方面广域网与局域网存在着比较大的区别。但是在网络层以及高层，它们可以使用共同的协议。

（3）如果仅简单地根据物理层和数据链路层的层次划分角度，将两种体系的内容放在一起讨论，对于初学者来说掌握起来有一定的困难。

（4）在本书的结构体系中：

第 3 章主要以广域网的物理层、数据链路层技术为背景讨论基于点-点通信线路的数据链路层协议与标准。

第 4 章以典型的 Ethernet 局域网与无线局域网 Wi-Fi 为背景，讨论基于广播信道的物理层和数据链路层协议与标准。

在第 3 章、第 4 章与第 5 章的基础上，第 5 章讨论它们可以共同使用的高层协议，即网络层的 IP 协议。这样的知识结构组织符合技术发展的规律，也能够适应读者循序渐进的学习需要。

第二部分　教学内容问答

问题 2-1：传输介质包括在物理层之中吗？

结论是：传输介质不包括在物理层之中。物理层处于 OSI 参考模型的最底层，它的任

务是实现二进制比特流的正确传输。计算机内部的二进制数字不能够直接通过传输介质传输。信号是数据在网络传输过程中电信号的表示形式。对于数据通信系统,它关心的是数据用什么样的电信号的形式来表示,如何去传输这些电信号。因此设置物理层的目的是:如何将计算机内部的二进制数据变换成传输介质可以传输的电信号;如何保证电信号传输的正确性。

因此,讨论物理层与物理传输介质的关系,需要注意以下几个问题。

(1) 物理层处于网络参考模型的最底层,它的上一层是数据链路层,它向下直接与传输介质相连。

(2) 物理层不是指与计算机相连接的具体的物理设备,或者具体的传输介质。物理层设计时主要考虑的是如何在连接开放系统的传输介质上传输各种数据的比特流。

(3) 由于计算机网络可以利用的物理传输介质与传输设备种类繁多,各种通信技术存在着很大的差异,而且各种新的通信技术又在快速发展,因此网络设计中,试图通过设置物理层,来尽可能地屏蔽这些差异,使数据链路层只需要考虑本层的服务与协议,而不需要考虑物理层具体使用了哪些传输介质与物理传输设备。

(4) 数据链路实体通过与物理层的接口,将数据传送给物理层,物理层按比特流的顺序,将信号传输到另一个物理层与数据链路实体。数据链路层在实现过程中不需要考虑物理传输设备与传输介质差异的存在。

(5) 由于计算机网络使用的通信线路分为两类:点-点通信线路和广播通信线路。点-点通信线路用于连接两个通信的结点;而广播通信线路的一条公共通信线路可以连接多个结点。广播通信线路又分为有线与无线两种。因此,物理层协议可以分为两类:基于点-点通信线路的物理层协议与基于广播通信线路的物理层协议。

问题2-2:为什么说物理层协议类型最复杂、变化最快?

早期流行的物理层协议标准是EIA-232-C标准。EIA-232-C标准是美国电子工业协会EIA在1969年制定的,它是基于点-点通信线路的串行、低速、模拟传输设备与计算机之间连接界面的物理接口标准。随着Internet接入技术的发展,每增加一种接入技术,就会增加一系列的物理层协议。例如,家庭接入主要通过ADSL调制解调器与电话线路接入,通过线缆调制解调器(Cable Modem)与有线电视同轴电缆接入。ADSL物理层协议定义了上行与下行传输速率标准、传输信号的编码格式与电平、同步方式、连接接口装置的物理尺寸等内容。Cable Modem有线电视电缆接入的物理层标准主要有“线缆数据业务接口规范”与IEEE 802.14的物理层标准,规定了线缆调制解调器的频带、上行与下行速率、信号调制方式与电平、同步方式等内容。不同的频带、上行与下行速率、信号调制方式与电平、同步方式,就会对应出现一种物理层协议。

在Ethernet的802.3标准中,从速率为10Mbps、100Mbps、1Gbps、10Gbps到40Gbps、100Gbps,每种速率的Ethernet的物理层就有多种协议。10Mbps以太网802.3标准的物理层协议包括10BASE-2、10BASE-5、10BASE-T;100Mbps快速以太网802.3u标准的物理层协议包括100BASE-T、100BASE-TX、100BASE-T4与100BASE-FX;千兆以太网802.3z标准的物理层协议包括1000BASE-T、1000BASE-CX、1000BASE-LX与1000BASE-SX;十千兆以太网的802.3ae标准的物理层协议包括局域网物理层协议LAN PHY标准与广域网物理层协议WAN PHY。

对于无线城域网 802.16 标准、无线局域网 802.11 标准与无线个人区域网 802.15.4 标准，根据所采用的覆盖范围、传输速率、通信频段、调制方式的不同，分别制定了多种物理层协议标准。因此，说物理层协议最复杂、变化最快是有道理的。

问题 2-3：如何理解数据通信中的同步方式问题？

同步技术是解决通信的收发双方在时间基准上的一致性问题，是数字通信中必须解决的一个重要问题。数据通信的同步包括以下两种：位同步、字符同步。

1. 位同步

实际计算机的时钟频率肯定存在着误差。时钟频率的积累误差可以造成接收比特取样周期的错误和传输数据的错误。因此，在数据通信过程中，首先要解决收发双方的时钟频率的一致性问题。实现位同步的方法主要有以下两种：外同步与内同步。

外同步法是在发送端发送一路数据信号的同时，另外发送一路同步时钟信号。内同步法则是从自含时钟编码的发送数据中提取同步时钟的方法。

2. 字符同步

保证收发双方正确传输字符的过程就叫作字符同步。实现字符同步的方法主要有以下两种：同步式与异步式。

同步传输将字符组织成组，以组为单位连续传送。每组字符之前加上一个或多个用于同步控制的同步字符 SYN，每个数据字符内不加附加位。异步传输是将每个字符作为一个独立的整体进行发送，字符之间的时间间隔可以是任意的。为了实现字符同步，每个字符的第一位前加起始位，字符的最后一位后加终止位。同步通信的传输效率要比异步通信的传输效率高，因此同步通信方式更适用于高速数据传输。

在讨论数据链路层时，还会遇到帧同步问题。

问题 2-4：传输介质特性需要从几个方面去描述？

传输介质是网络中连接收发双方的物理通路，也是通信中实际传送数据信号的载体。网络中常用的传输介质有：双绞线、同轴电缆、光纤与无线通信信道。

研究传输介质需要注意了解传输介质特性。传输介质的特性对网络中数据通信质量的影响很大，这些特性主要包括如下几个。

(1) 物理特性：对传输介质物理结构的描述。

(2) 传输特性：传输介质允许传送数字或模拟信号，以及调制技术、传输容量与传输的频率范围。

(3) 连通特性：允许点-点或多点连接。

(4) 地理范围：传输介质的最大传输距离。

(5) 抗干扰性：传输介质防止噪声与电磁干扰对传输数据影响的能力。

(6) 相对价格：包括器件、安装与维护费用。

对于学习计算机网络知识与实际组网的能力，需要学生重点掌握双绞线、光纤与无线通信信道的知识，同轴电缆目前使用得比较少了。但是对于理解传统 Ethernet，了解同轴电缆的特性还是有必要的。

问题 2-5：学习双绞线知识需要注意哪些问题？

1. 双绞线的概念

双绞线(Twisted Pair，TP)是一种综合布线工程中最常用的传输介质，是由两根具有绝

缘保护层的铜导线组成的。把两根绝缘的铜导线按一定密度与角度互相绞在一起,使得导线之间的电磁波降低到最小。双绞线过去主要是用来传输模拟信号的,现在同样可以传输数字信号。双绞线是由多对双绞线一起包在一个绝缘电缆套管里,构成了双绞线电缆。实际上人们一般是将“双绞线电缆”简称为“双绞线”。局域网中所使用的双绞线分为两类:屏蔽双绞线(Shielded Twisted Pair,STP)与非屏蔽双绞线(Unshielded Twisted Pair,UTP)。

2. 双绞线的类型

常见双绞线可以分为:3类线、5类线和超5类线,以及最新的6类线与7类线。

(1) 3类线(CAT3):3类线最高传输速率为10Mbps,主要应用于语音与10Mbps的Ethernet(10BASE-T),采用RJ-45连接器。

(2) 5类线(CAT5):5类线增加了绕线密度,外套一种高质量的绝缘材料,线缆最高频率带宽为100MHz,最高传输率为100Mbps,用于语音传输和最高传输速率为100Mbps的数据传输,主要用于100BASE-T和1000BASE-T网络,最大网段长为100m,采用RJ-45连接器。

(3) 超5类线(CAT5e):超5类具有衰减小、串扰少、延时小的特点,并且具有更高的衰减与串扰的比值(ACR)和信噪比,主要用于1Gbps的千兆位以太网中。

(4) 6类线(CAT6):6类线的传输频率为1~250MHz,能够提供二倍的超5类的带宽。6类线的传输性能远远高于超5类标准,最适用于传输速率高于1Gbps的应用。

(5) 超6类或6A(CAT6A):超6类线传输带宽介于6类和7类之间,传输频率为500MHz。

(6) 7类线(CAT7):带宽为600MHz,可能用于10Gbps的Ethernet中。

需要注意的是:目前国家还没有出台正式6类线与7类线产品的检测标准,只能采用由各个厂家公布的测试参数。

问题2-6:学习光纤物理层标准需要注意哪些问题?

对于物理层标准,早期比较成熟和广泛应用的是在电话线路上用调制解调器Modem的串行通信的物理层标准——RS-232C,因此一般的教科书多是分析这个协议。随着光纤通信、光纤入户与光以太网的广泛应用,关于光纤通信的物理层标准逐渐变得越来越重要。目前广域网、城域网基本上都是采用“路由器+光纤”的组网模式。接入技术中,光纤到家、光纤到楼、光纤到路边、光纤到结点、光纤到办公室非常普遍。但是,教科书中很少系统地讨论这个问题,对于光纤物理层标准知识的信息,需要注意以下几个方面的问题。

1. 对于光纤特点的了解

光纤的特点主要表现在以下几个方面。

(1) 由于光纤具有低损耗、宽频带、高数据传输速率、低误码率与安全保密性好的特点,因此是一种最有前途的传输介质。

(2) 光纤传输的类型分为单模与多模两类。单模光纤的性能要优于多模光纤。

(3) 光缆一般由三部分构成:缆芯、中心加强芯与护套。目前,光缆在广域网、城域网与局域网,以及在电信传输网、广播电视传输网中都得到广泛的应用。

(4) 由于光纤传输速率高、误码率低、安全性好,因此光纤已经成为计算机网络中最有发展前景的传输介质。同时,由于光纤通信技术的发展,光纤组网成本的降低,光纤已经从主要用于连接广域网核心路由器,逐渐发展到城域网、局域网,目前正在向光纤直接接入办

公室、光纤接入家庭的方向发展。

2. 光纤物理层标准

随着光纤应用范围的扩大，很多终端用户已经开始接触到光纤，也会接触到物理层关于光纤的传输速率、传输距离等参数的问题。例如，高速 Ethernet 的物理层就制定了多个关于光纤的物理层标准，其中涉及多个描述物理层特征的参数。了解有关光纤的物理层标准，需要注意以下几个问题。

(1) 影响光纤传输距离的因素主要有传输模式、光载波的频率、光纤的尺寸。

(2) 计算机产生的电信号需要在传输时变换成光载波信号在光纤上传播。由于光纤只能够单方向传输光载波信号，因此要实现计算机与交换机的双向传输需要使用两根光纤。

(3) 在物理层协议中，用于从计算机向交换机传送信号的光纤称为上行光纤，用于从交换机向计算机传送信号的光纤称为下行光纤。上行光纤与下行光纤使用不同的光载波频率。

(4) 物理层协议规定的物理参数主要包括传输模式、上行光纤与下行光纤光载波的频率、光纤的尺寸、光接口，以及最大光纤传输距离。

例如，在传输速率为 1Gbps 的千兆以太网 GE 的物理层 1000BASE-LX 标准中，规定：传输介质采用单模光纤，光纤直径大于 10μm，上行光纤与下行光纤的光载波的频率分别为 1270 nm 与 1355nm，光纤最大长度为 5km。

问题 2-7：什么是"裸光纤"？

我们在讨论网络通信问题时会遇到"裸光纤"这个术语。对于"裸光纤"术语的理解需要注意以下几点。

(1) 光纤制造商将刚拉制出来而尚未进行一次涂覆的光纤称为裸光纤。纤芯的直径约为 50μm，外部包层直径约为 125～130μm，纤芯涂覆了外部包层之后就形成了能够稳定传输光载波的光纤通道。

(2) 对于电信运营商来说，如果两个结点之间的距离超过 3km，一般就需要用光纤来连接。电信运营商一般都提供裸光纤租用业务，用户可以通过租用运营商的裸光纤来组建大型企业、校园网、园区网的宽带主干网。在这种情况下，裸光纤可以理解为中间没有接续设备的光纤。

问题 2-8：学习无线通信知识需要注意哪些问题？

随着移动互联网、无线城市的应用，学习无线通信知识也就越来越重要了。学习无线通信知识需要注意以下几个问题。

1. 无线通信的基本知识

描述电磁波的参数有三个：波长、频率与光速。它们三者之间的关系为：

$$\lambda \times f = C$$

其中，光速 C 为 3×10^8m/s，频率 f 的单位为 Hz。

电磁波的传播有两种方式：一种是在自由空间中传播，即通过无线方式传播；另一种是在有限制的空间区域内传播，即通过有线方式传播。用同轴电缆、双绞线、光纤传输电磁波的方式属于有线方式传播。在同轴电缆中，电磁波传播的速度大约等于光速的 2/3。很多例题中取电磁波传播的速度为 3×10^8m/s 就是这样来的。同时需要注意的是，这是指电磁

波在同轴电缆等有线传输介质中的传播速度。

2. 无线信号频率、功率与覆盖范围

理解无线信号频率、功率与覆盖范围的概念需要注意以下几个问题。

(1) 图2-2给出了无线通信的示意图。在自由空间中,结点A发射信号功率为E,信号经过距离为r的传播后到达结点B的位置,这个位置的信号功率为E/r^2。

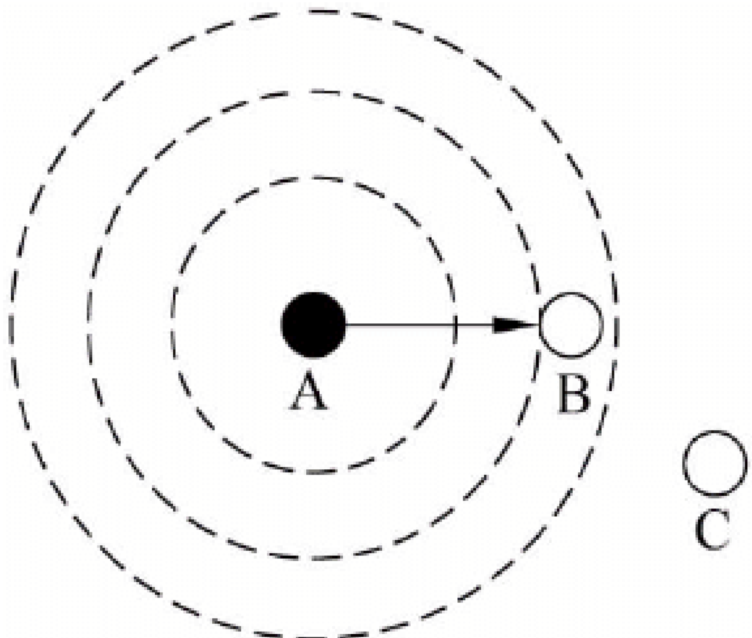


图2-2 无线通信覆盖范围示意图

(2) 无线通信中,接收结点必须处于发送结点无线电波覆盖的范围之内,并且实际信号接收功率要大于接收设备最小接收功率的限制。

(3) 无线通信中,描述无线信号的参数主要是:频率与信号强度。接收主机通过接收机接收无线信号有两个基本条件:一是发送信号频率要在接收机的频率范围之内;二是接收到的信号强度要大于接收机的接收灵敏度。例如,主机B与主机C的接收机频带为2.45～2.48GHz,主机A发送的信号频率为2.465GHz,处于主机B与C接收信号频带之内,满足第一个基本条件。接收机B、C的接收灵敏度都为－60dBm,接收机B接收到的无线信号强度为－50dBm,大于接收机的接收灵敏度;而接收机C接收到的无线信号强度为－70dBm,小于接收机C的接收灵敏度。那么主机B的接收机能够接收主机A发送的无线信号,而主机C的接收机不能够接收主机A发送的无线信号。这样我们可以说:主机B处于主机A的信号覆盖范围之内,主机C不处于主机A的信号覆盖范围之内。

(4) 这里所说的信号强度是指信号功率。信号功率单位是瓦(W)或毫瓦(mW)。在无线局域网802.11协议的讨论中,通常使用的是信号功率的相对值,即dBm。dBm是指信号功率相对于1mW的dB值。计算公式为:$dBm=10\times\log_{10}(P_{mW})$,其中,$P_{mW}$是信号以mW为单位的功率值。表2-1给出dBm与$P_{mW}$的对照表。

表2-1 dBm与P_{mW}对照表

dBm	P_{mW}	dBm	P_{mW}
＋20dBm	100mW	－40dBm	0.000 1mW
＋10dBm	10mW	－50dBm	0.000 01mW
0dBm	1mW	－60dBm	0.000 001mW
－10dBm	0.1mW	－70dBm	0.000 000 1mW
－20dBm	0.01mW	－80dBm	0.000 000 01mW
－30dBm	0.001mW		

从表2-1中可以看出,1mW是一个参考点,0dBm表示1mW。如果测量值是＋dBm,表示信号强度大于1mW;如果测量值是－dBm,表示信号强度小于1mW。大部分802.11无线信号发射功率一般在100mW内,可以表示为＋20dBm,而无线网卡接收到的信号功率一般只有0.0001mW,可以表示为－40dBm。由于距离增加与其他因素引起信号强度衰减,接收信号功率仅为0.000 000 000 1mW,即－100dBm是常见的事,显然用－100dBm表示是一个非常简洁和不容易出错的方法。在802.11网络现场勘测中,使用的信号强度测量仪器也以dBm为单位来记录不同地理位置的无线信号强度。

问题 2-9：什么是工业、科学与医药专用 ISM 频段？

为了维护无线通信的有序性，防止不同通信系统之间的干扰，世界各国都要求无线电频段的使用者向政府管理部门申请特定的频段，获得批准后才可以使用。

同时，国际电信联盟无线通信局 ITU-R 要求世界各国专门划出免予申请的工业、科学与医药的 ISM 频段(Industrial Scientific Medical Band)，即专门开放某一些频段给工业、科学和医学机构使用。原则上使用这些频段的用户不需要事先申请许可证，也不需要缴纳费用，只需要遵守一定的发射功率(一般低于 1W)限制，并且不要对其他频段造成干扰即可。

ISM 频段在各国的规定并不统一。例如，在美国有三个频段 902～928MHz、2400～2484.5 MHz 及 5725～5850MHz，而在欧洲 900MHz 的频段有一部分用于 GSM 通信，而 2.4GHz 是世界各国共同的 ISM 频段。因此无线局域网(IEEE 802.11b/IEEE 802.11g)与蓝牙、ZigBee 等无线网络均可工作在 2.4GHz 频段上。

ITU-R 指定的 ISM 频段的频率分配如表 2-2 所示。

表 2-2　ISM 频带分配

频率范围	中心频率	可用性
6.765～6.795MHz	6.780MHz	取决于本地
13.553～13.567MHz	13.560MHz	
26.957～27.283MHz	27.120MHz	
40.66～40.70MHz	40.68MHz	
433.05～434.79MHz	433.92MHz	
902～928MHz	915MHz Region 2 only	
2.420～2.4835GHz	2.450GHz	
5.725～5.875GHz	5.800GHz	
24～24.25GHz	24.125GHz	
61～61.5GHz	61.25GHz	取决于本地
122～123GHz	122.5GHz	取决于本地
244～246GHz	245GHz	

问题 2-10：如何认识 CDMA 与 OFDM？

认识这个问题，需要注意以下几点。

1. 无线网络的物理层

随着无线网络技术的发展，人们开始在计算机网络物理层实现技术中注意到 CDMA 与 OFDM 技术的问题。有的教材试着讨论 CDMA 的特殊码型设计问题。实际上，在计算机网络的无线局域网(IEEE 802.11)或无线城域网(IEEE 802.16)中都没有使用到码分多址(CDMA)技术，而是使用正交频分复用(OFDM)技术。本教材没有对码分多址(CDMA)技术展开讨论，同时 OFDM 技术十分复杂，对于读者掌握计算机网络的基本工作原理并不是至关重要的，因此也没有展开讨论。对于教师需要有一定的背景知识。

2. 码分多址技术

码分多址 CDMA 实际上是第二次世界大战时期为了适应军事通信抗干扰而研究的一种直接序列扩频通信技术，1995 年第一个商用的基于 CDMA 的蜂窝移动通信系统正式使用，目前已经成为 3G 的重要技术标准之一。

CDMA 的特点主要表现在以下两个方面。

(1) CDMA 技术支持多个移动通信用户在同一个时间,使用相同的频段进行通信;各个用户使用特殊设计的码型,它们之间不会互相干扰。

(2) CDMA 通信系统为每个用户指派一种唯一的码型。它将每个比特时间划分为 m 个码片;m 值一般为 64 或 128;由 m 个码片组成 m 比特的码片序列。如果一个用户需要发送 100 比特的数据,那么 CDMA 通信系统实际需要发送的码片序列为 $100 \times m$ 比特。如果一个用户要求的发送速率为 S(bps),那么 CDMA 系统实际发送的速率应该为 $m \times S$(bps)。所以,CDMA 属于扩频通信中的直接序列扩频 DSSS 中的一种。

作为 3G 的主流技术,CDMA 有三种主要的协议标准:国际电信联盟 ITU 的宽带码分多址 W-CDMA 标准、美国高通(Qualcomm)公司开发的 CDMA2000 标准,以及我国有自主知识产权的时分同步码分多址(Time Division-Synchronous CDMA,TD-SCDMA)标准。

3. 正交频分复用技术

正交频分复用 OFDM 技术是在频分多路复用 FDM 基础上发展起来的。OFDM 是多载波调制(Multi-Carrier Modulation,MCM)方法的一种。OFDM 是将信道分成若干个正交的子信道,将高速数据信号转换成 n 个并行的低速子数据流,调制到在每个子信道上进行传输。正交信号可以在接收端采用相关技术来分开,实现频分多路复用。

OFDM 已经广泛应用于各种数字通信中,如移动无线调频通信系统、高比特率数字用户线(HDSL)系统、非对称数字用户线(ADSL)系统、甚高比特率数字用户线(VDSL)系统。IEEE 802.11a 无线局域网的物理层采用 OFDM 调制方法,使得传输速率可以达 54Mbps。无线城域网标准 IEEE 802.16 在 2~11GHz 频段也采用了 OFDM 调制方法。

问题 2-11:如何理解信息、数据和信号之间的关系?

图 2-3 从人与人通过计算机网络交互的过程看信息、数据和信号关系的示意图。

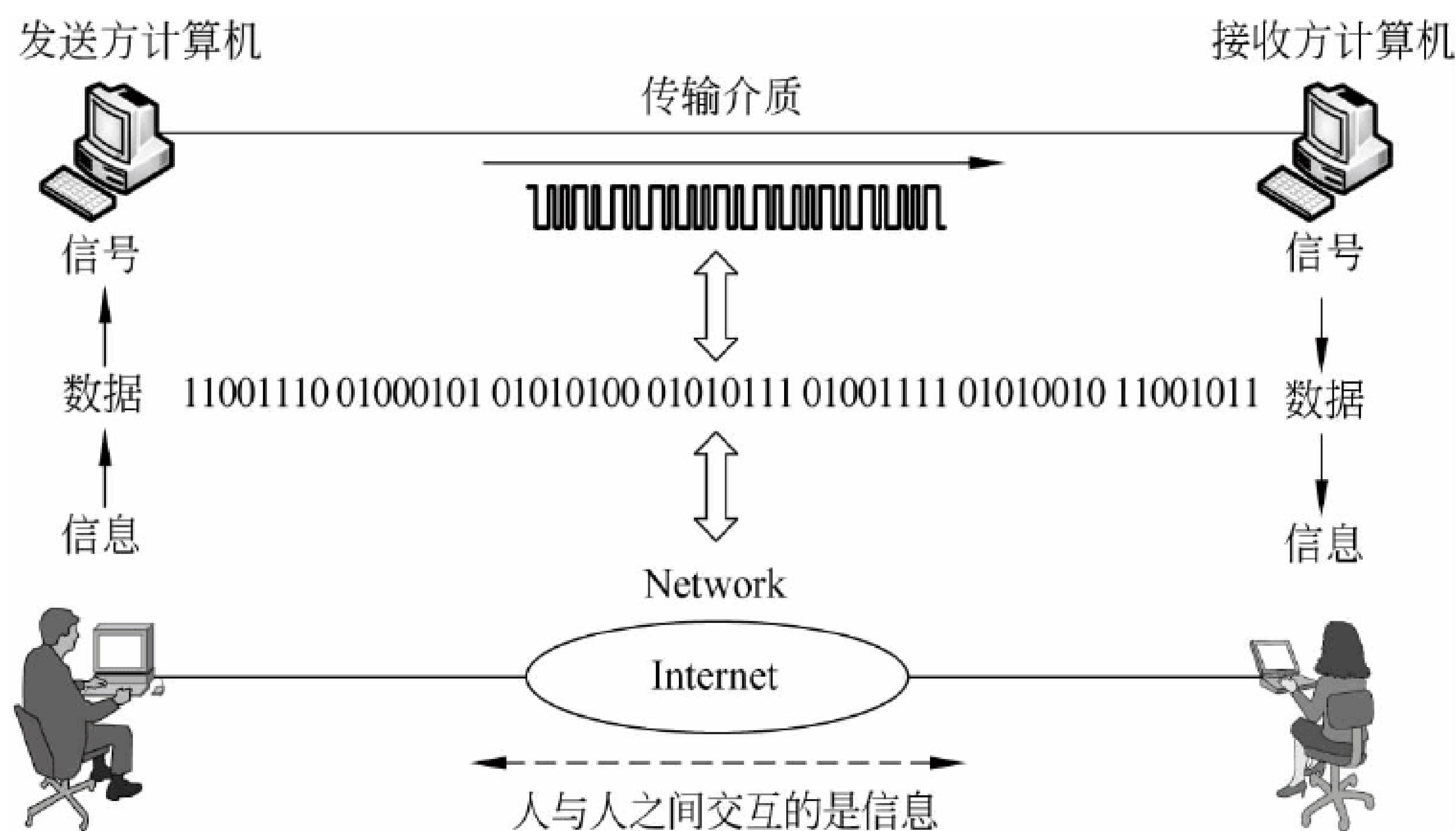


图 2-3 信息、数据和信号之间的关系

(1) 人与人交互的是信息。作为计算机网络的使用者,用户关心的是通过网络获取的信息。信息的载体可以是文字、语音、图形、图像或视频。

(2) 人用手工的方式将信息输入到计算机,计算机将输入的文字、语音、图形、图像或视频信息由特定的编码方式变换成二进制数据,储存到计算机之中,按照用户的需求进行

处理。

(3) 如果需要在不同的网络用户之间进行信息交互，就需要将计算机内部的二进制数据变换成可以在传输介质上传输的信号。

问题 2-12：数据编码分类的依据是什么？

这是理解物理层很多重要概念的基础。可以从以下两个方面来回答这个问题。

(1) 计算机中的数据是以离散的二进制 0、1 比特序列方式表示的。计算机数据在传输过程中的数据编码类型，主要取决于它采用的通信信道所支持的数据通信类型。

(2) 根据数据通信类型来划分，网络中常用的通信信道分为两类：模拟通信信道与数字通信信道。相应地，用于数据通信的数据编码方式也分为两类：模拟数据编码与数字数据编码。网络中数字数据编码的方案是很多的，并且随着高速网络技术的发展，已经出现了一系列的新技术，但是最基本的数据编码方式可以归纳为图 2-4 给出的几种类型。

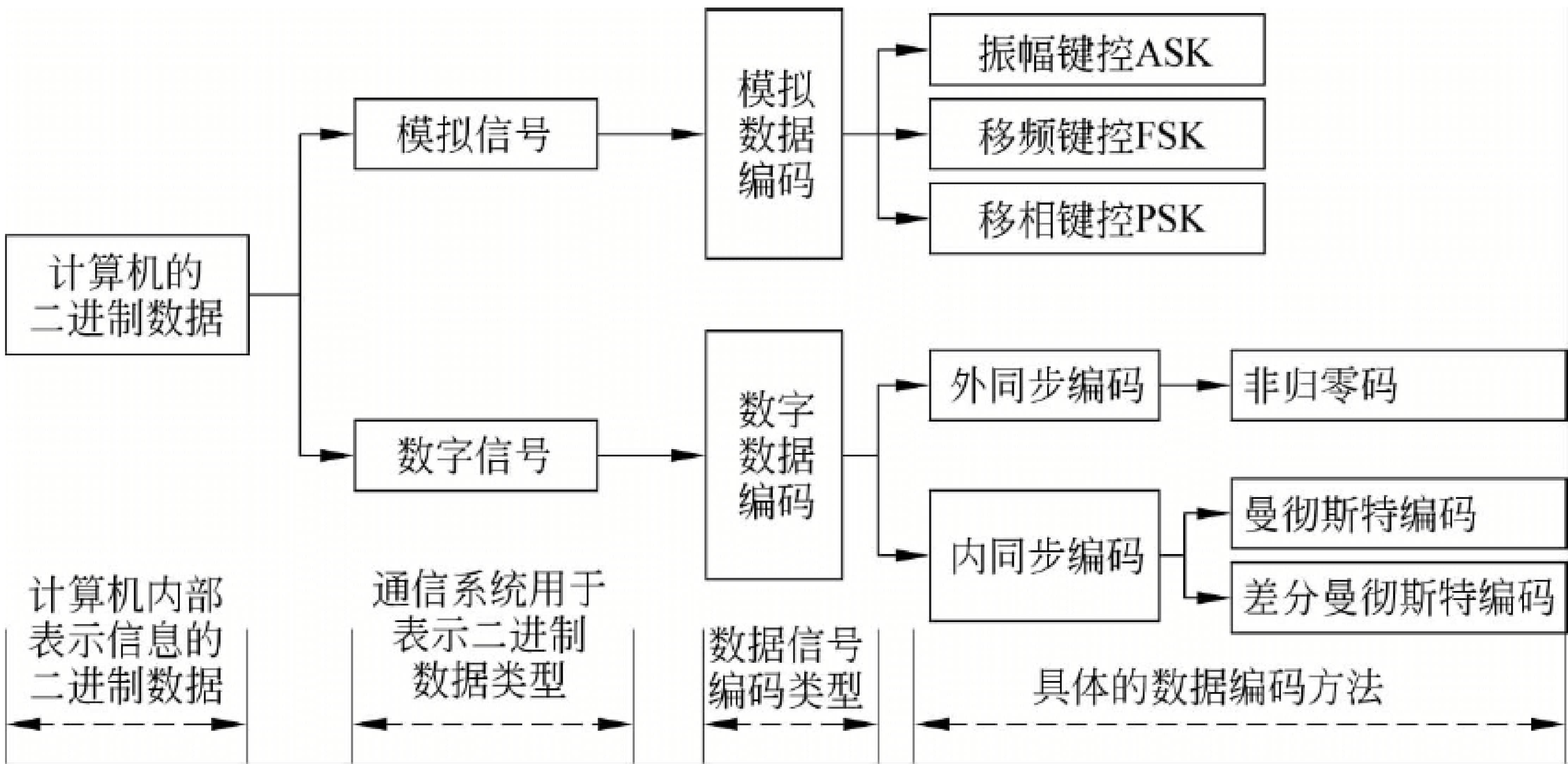


图 2-4　数据编码类型

问题 2-13：如何理解频带传输与模拟数据信号编码方法？

电话线路是典型的模拟通信线路，它是目前世界上覆盖面最广、应用最广泛的通信线路。无论网络与通信技术如何发展，电话仍是一种基本的通信手段。传统的电话线路是为传输语音信号而设计，只适用于传输音频范围(300～3400Hz)的模拟信号，无法直接传输计算机的二进制数字信号。为了利用模拟语音通信的电话交换网实现计算机的数字数据信号的传输，必须首先将数字信号转换成模拟信号。利用模拟通信线路传输数据信号的方法称为频带传输。

在模拟数据信号编码形成过程中，首先选择音频范围内的某一角频率 ω 的正(余)弦信号作为载波，该正(余)弦信号可以写为：$u(t)=u_m \cdot \sin(\omega t+\varphi_0)$。在载波 $u(t)$ 中，有三个可以改变的电参量(振幅 u_m、角频率 ω 与相位 φ)。可以通过变化三个电参量，来实现模拟数据信号的编码。图 2-5 给出了模拟数据信号的编码方法示意图。

1. 振幅键控

振幅键控(Amplitude Shift Keying，ASK)方法是通过改变载波信号振幅来表示数字信号 1、0。例如，可以用载波幅度为 u_m 表示数字 1，用载波幅度为 0 表示数字 0。图 2-5(a)给出了 ASK 信号波形，其数学表达式为：

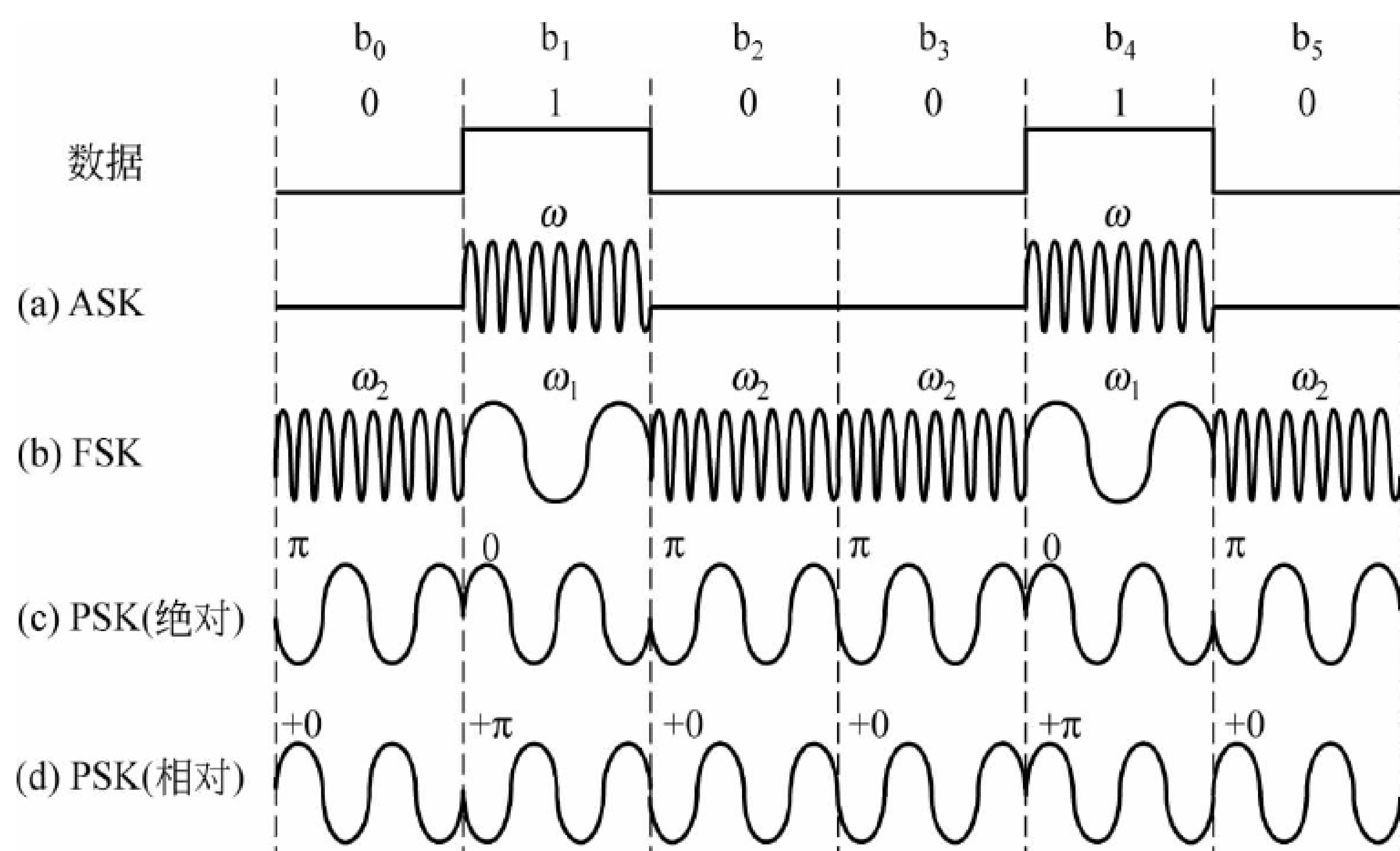


图 2-5 模拟数据信号的编码方法

$$u(t)=\begin{cases}u_m\cdot\sin(\omega_1 t+\varphi_0) & \text{数字 1}\\ 0 & \text{数字 0}\end{cases}$$

ASK 信号实现容易、技术简单,但是抗干扰能力较差。

2. 移频键控

移频键控(Frequency Shift Keying,FSK)方法是通过改变载波信号角频率来表示数字信号 1、0。例如,可以用角频率 ω_1 表示数字 1,用角频率 ω_2 表示数字 0。图 2-5(b)给出了 FSK 信号波形,其数学表达式为:

$$u(t)=\begin{cases}u_m\cdot\sin(\omega_1 t+\varphi_0) & \text{数字 1}\\ u_m\cdot\sin(\omega_2 t+\varphi_0) & \text{数字 0}\end{cases}$$

FSK 信号实现容易、技术简单,抗干扰能力较强,是目前最常用的调制方法之一。

3. 移相键控

移相键控(Phase Shift Keying,PSK)方法是通过改变载波信号的相位值来表示数字信号 1、0。如果用相位的绝对值表示数字信号 1、0,则称为绝对调相。如果用相位的相对偏移值表示数字信号 1、0,则称为相对调相。

在载波信号 $u(t)$ 中,φ_0 为载波信号的相位。最简单的情况是:用相位的绝对值来表示它所对应的数字信号。图 2-5(c)给出了绝对调相的信号波形。当表示数字 1 时,取 $\varphi_0=0$;当表示数字 0 时,取 $\varphi_0=\pi$。这种简单的绝对调相方法可以用下式表示:

$$u(t)=\begin{cases}u_m\cdot\sin(\omega t+0) & \text{数字 1}\\ u_m\cdot\sin(\omega t+\pi) & \text{数字 0}\end{cases}$$

相对调相用载波在两位数字信号的交接处产生的相位偏移来表示载波所表示的数字信号。最简单的相对调相方法是:两比特信号交接处遇 0,载波信号相位不变;两比特信号交接处遇 1,载波信号相位偏移 π。图 2-5(d)给出了相对调相的信号波形。

在实际使用中,移相键控方法可以方便地采用多相调制方法达到高速传输的目的。移相键控方法的抗干扰能力强,但是实现技术比较复杂。

以上讨论的是二相调制的方法,用两个相位值分别表示二进制数 0、1。在模拟数据通信中,为了提高数据传输速率,人们常采用多相调制的方法,称为正交相移键控(Quadrature

Phase Shift Keying,QPSK)。例如,将待发送的数字信号按两比特一组的方式组织,两位二进制比特可以有4种组合,即00、01、10、11。每组是一个双比特码元,用4个不同相位值表示这4种双比特的码元。在调相信号传输过程中,相位每改变一次,传送两个二进制比特。这种调相方法称为四相调制。同理,如果将发送的数据每三个比特组成一个三比特码元组,三位二进制数共有8种组合,对应可以用8种不同的相位值表示,这种调相方法称为八相调制。在多相调制中,相位每改变一次产生一个码元;一个码元可以传送两个或三个二进制比特。

问题2-14:如何理解调制解调器的基本工作原理?

尽管目前调制解调器(Modem)已经不使用,但是调制解调的概念在很多地方都会涉及,因此作为教师需要了解一些调制解调器的知识。

1. 调制解调器的基本工作原理

在频带传输中,计算机通过Modem与电话线路连接。在发送端,Modem将计算机产生的数字信号转换成电话交换网可以传送的模拟数据信号(如ASK、FSK或PSK方式);在接收端,Modem将接收到的模拟数据信号还原成数字信号传送给计算机。在全双工通信方式中,Modem应具有同时发送与接收模拟数据信号的能力。计算机通过Modem与电话交换网实现远程通信的结构如图2-6所示。

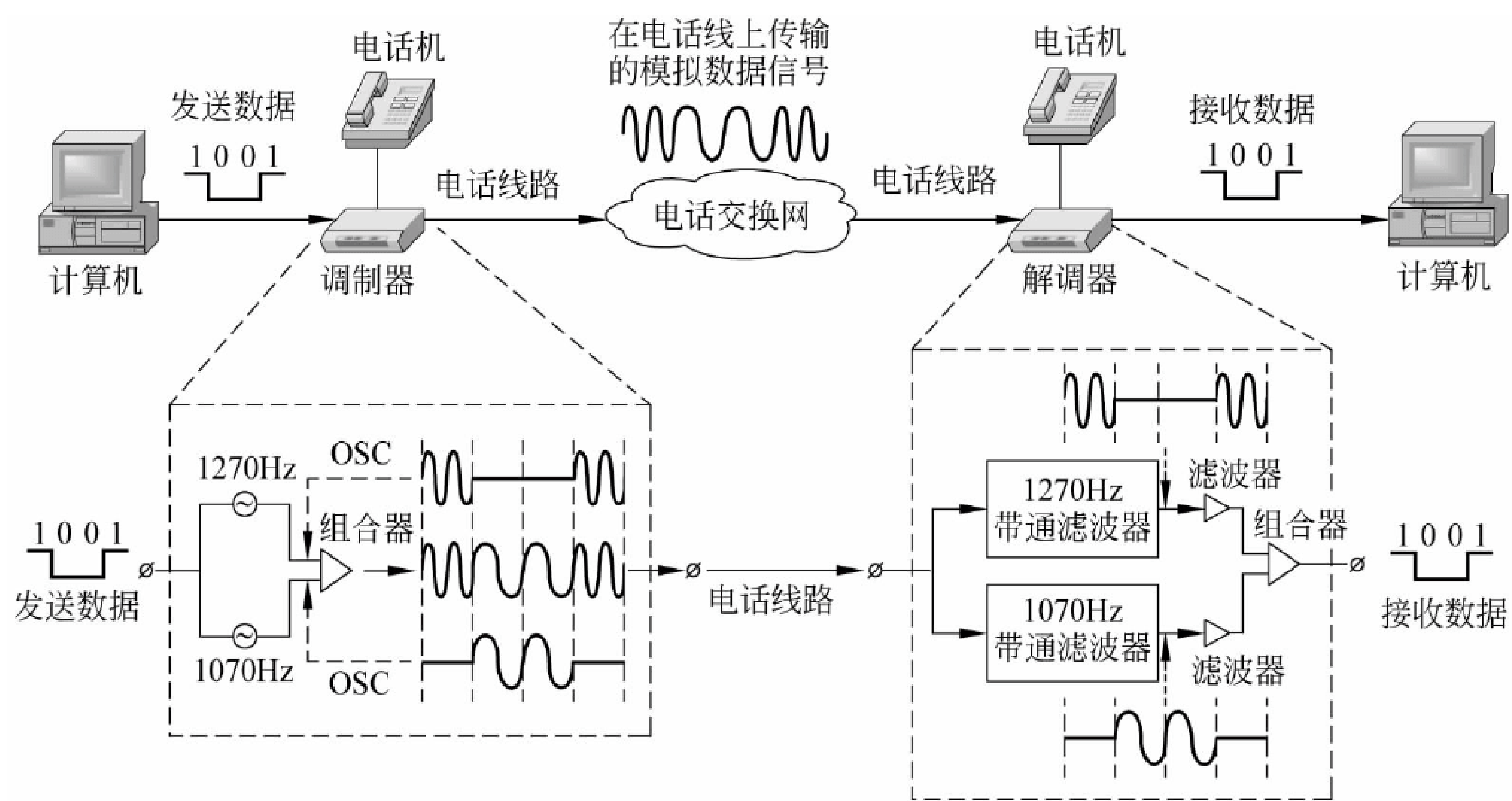


图2-6 实现FSK方式的Modem工作原理示意图

图2-6以移频键控FSK方式为例,给出了Modem工作原理示意图。理解Modem工作原理需要注意以下几个问题。

(1) 发送方调制器是用输入的数字脉冲信号控制两个不同频率振荡器信号的输出,来实现数字信号-模拟信号的转换。当输入的数字脉冲信号为高电平(对应于逻辑1)时,频率f_1=1270Hz的振荡器有信号输出,当输入的数字脉冲信号为低电平(对应于逻辑0)时,频率f_2=1070Hz的振荡器有信号输出。

(2) 在调制器的输出端,按照输入的数字脉冲信号1、0序列排列顺序控制的两种频率的正(余)弦信号通过组合器组合起来,构成了移频键控FSK信号。

(3) 由于对应1、0的两种不同频率的正(余)弦信号是处于电话交换网的通频带内，因此模拟数据信号FSK可以顺利地通过模拟电话交换线路传送到接收方。

(4) 在接收方通过设置对应 f_1、f_2 两种频率的带通滤波器，将两种不同频率的正(余)弦信号分开，使频率为 f_1 和 f_2 的正(余)弦信号分别通过两个检波器，再将检波器输出信号送给组合器叠加。组合器输出的解调信号对应的数字脉冲信号的高、低电平(即逻辑1与0)的变化规律，与调制器输入的数字数据信号的高、低电平变化规律相同，因而能正确地还原数字数据信号。

2. 调制解调器的全双工通信实现方法

在完成调制、解调工作原理的初步讨论后，进而要讨论Modem如何实现在一对电话线上完成全双工通信的工作原理。图2-7给出Modem实现全双工通信的工作原理。

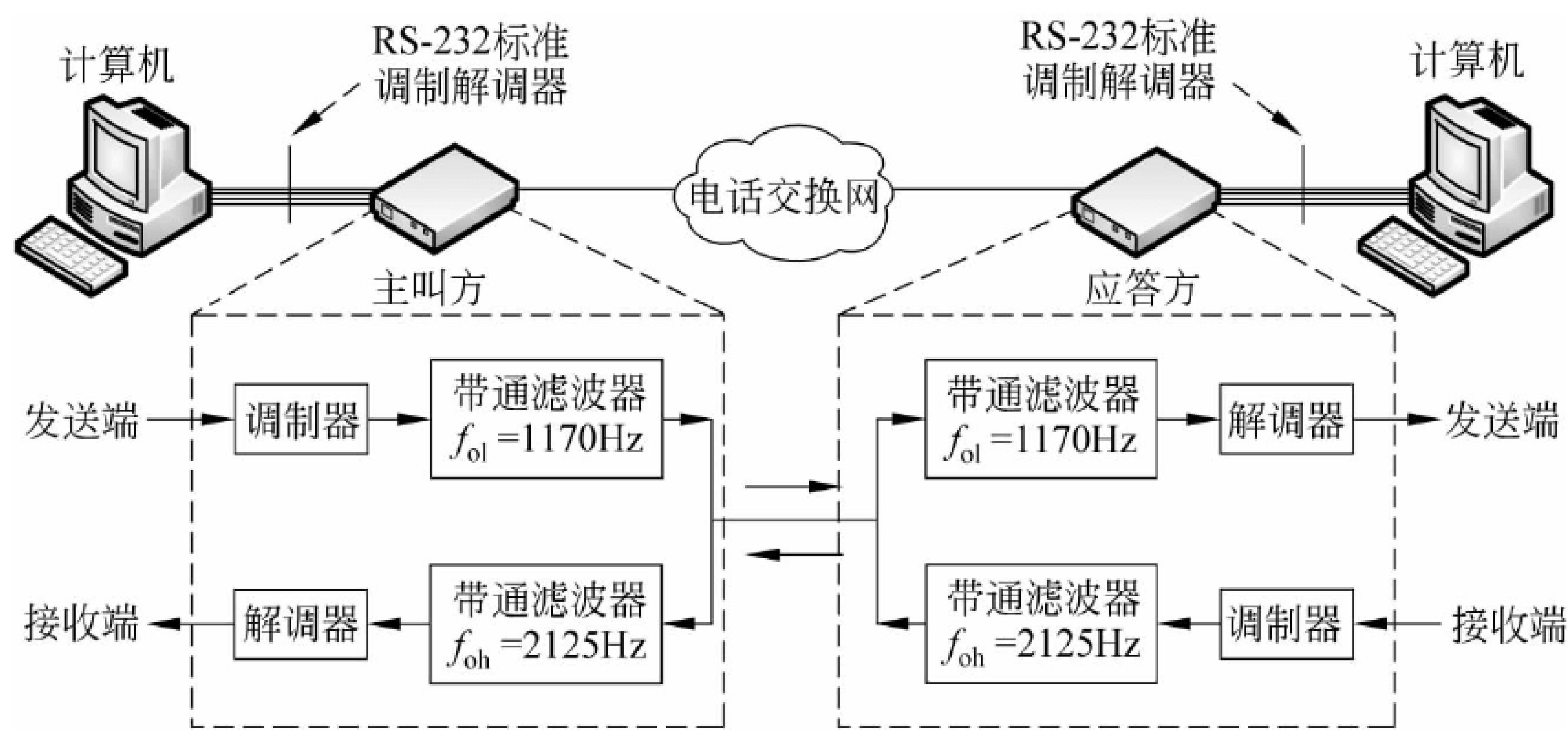


图2-7　Modem实现全双工通信的工作原理

理解Modem全双工通信工作原理，需要注意以下几个问题。

(1) 在实际计算机通信中，任何一台计算机都需要同时具备发送和接收数据的能力。为了实现在一对电话线上全双工通信，标准的Modem都规定了两个频率值，即上、下频带。

(2) 在一次数据通信中，主动发起通信的一端叫作呼叫方，被动参加通信的一端叫作应答方。通信的两台计算机调制解调器中谁是呼叫方，谁是应答方，完全根据在一次通信过程中是主动发起通信，还是被动地响应通信来动态决定的，而不是固定的。呼叫端与应答端的关系如图2-8所示。

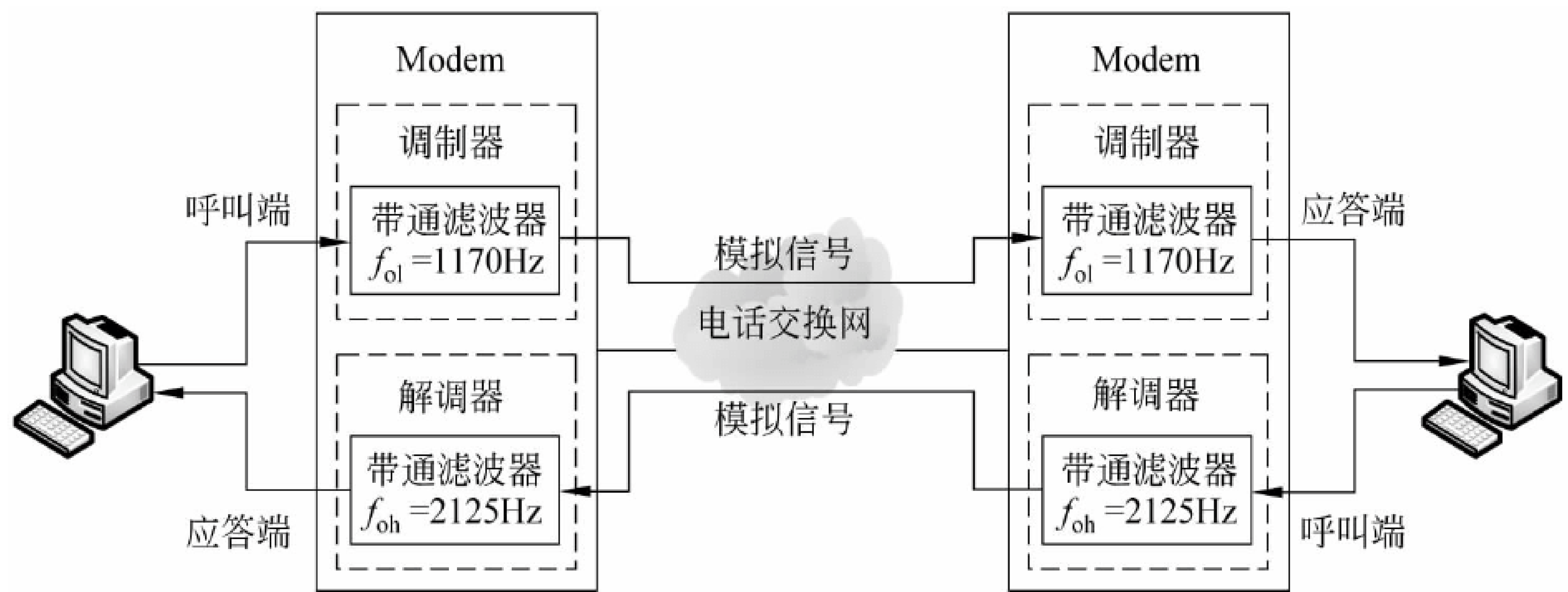


图2-8　呼叫端与应答端的关系

（3）物理层协议 RS-232-C 规定：如果一方被确定为呼叫方，则它使用下频带发送数据，在上频带接收数据。那么另一方一定是应答方，它在发送数据时使用上频带，接收时使用下频带。

（4）同时，RS-232-C 规定：FSK Modem 使用的上、下频带对应于数字“1”“0”的信号频率如下。

上频带　数字 1→2225Hz
　　　　数字 0→2025Hz
下频带　数字 1→1270Hz
　　　　数字 0→1070Hz

上频带的中心频率为 2125Hz，下频带的中心频率为 1170Hz。为了实现在一对电话线上进行全双工通信，Modem 通过对应于上频带中心频率 2125Hz 与下频带中心频率 1170Hz 的两种带通滤波器，将双方的发送与接收通道分开，达到全双工通信的目的。

（5）带通滤波器是一种选频电路，它只允许频率为 $f_0 \pm \Delta f$ 的范围内的信号通过，$2\Delta f$ 为带通滤波器的通频带。f_0 与 Δf 可以在带通滤波器电路设计中根据需要加以设定。调制解调器的下频带所用的低端带通滤波器中心频率 f_{01} 为 1170Hz，Δf 为 200Hz，那么它只允许频率为 970～1370Hz 的信号通过，即 1070Hz 与 1270Hz 的信号可以通过。同理，上频带所使用的高端带通滤波器，也只允许频率为 2025Hz 与 2225Hz 的信号通过。上、下频带的频率分配情况如图 2-9 所示。

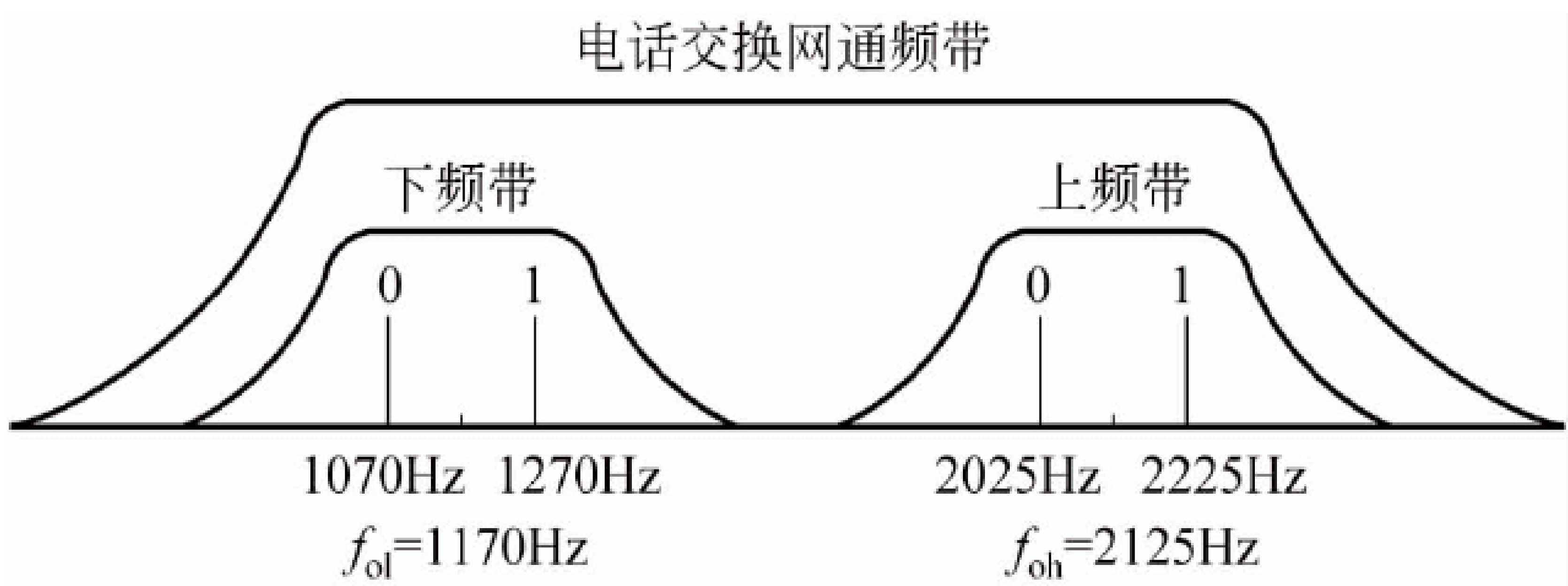


图 2-9　FSK 的上、下频带的频率分配

问题 2-15：EIA RS-232 物理接口标准包括哪些基本的内容？

与调制解调器问题一样，尽管目前很少用到 RS-232 物理接口标准了，但是这个概念在很多地方都会涉及，因此作为教师需要了解一些 RS-232 物理接口标准的知识。了解 RS-232 物理接口标准，需要注意以下几点。

（1）EIA-232-C 标准是美国电子工业协会在 1969 年制定的一种通信标准，它是基于点-点通信线路、串行、低速线路的模拟传输设备与计算机之间连接界面的物理接口标准。EIA-232-C 标准规定了计算机串行通信接口卡与调制解调器 Modem 之间物理接口的机械、电气、功能和规程的具体参数与工作流程。目前，很多低速的数据通信设备仍然采用 RS-232-C 标准，以及在此基础上发展的 EIA RS-449 与 CCITT X. 21 标准等。

（2）物理层标准与物理接口标准是有区别的。OSI 参考模型中物理层标准化工作要比数据链路层、网络层等高层慢。形成这种情况的原因主要有两点。一是与物理层涉及的具体物理设备、传输介质与通信手段的复杂性有关；另一个更重要的原因是在 ISO 提出 OSI 参考模型之前，许多属于物理层的模型和协议就已经提出，并在某些领域已形成相当的工业

生产规模和广泛的应用。这些模型、协议没有严格遵循分层的方法与原则,也没有像 OSI 那样分为服务定义与协议的规则说明。在现实情况下,要把已有的物理层模型和协议统一到 OSI 物理层服务定义与协议说明的框架之下难度是很大的。目前使用的物理层协议是物理接口标准,这种物理接口标准定义了物理层与物理传输介质之间的接口。最常用的物理接口标准是 EIA-232-D、EIA RS-449 与 CCITT X.21。

(3) 物理接口的主要特性。

物理接口标准的主要特性包括机械特性、电气特性、功能特性与规程特性。

① 机械特性。

机械特性规定了物理连接时所使用的可接插连接器的形状和尺寸,连接器中引脚的数量与排列情况等。

② 电气特性。

电气特性规定了在物理连接上传输二进制比特流时,线路上信号电平高低、阻抗及阻抗匹配、传输速率与距离限制。早期的标准定义了物理连接边界点上的电气特性,而较新的标准定义了发送器和接收器的电气特性,同时给出互连电缆的有关规定。新的标准更利于发送和接收电路的集成化工作。

③ 功能特性。

功能特性规定了物理接口上各条信号线的功能分配和确切定义。物理接口信号线一般分为:数据线、控制线、定时线和地线等几类。

④ 规程特性。

规程特性定义了利用信号线进行二进制比特流传输的一组操作过程,包括各信号线的工作规则和时序。

问题 2-16:如何理解波特率与比特率的定义?

早期在模拟线路上使用调制解调器进行数据通信时使用过调制速率与波特率的概念。理解波特率的概念需要注意以下几个问题。

1. 波特率的定义

调制速率描述通过模拟线路传输模拟数据信号的过程中,从调制解调器输出的调制信号每秒钟载波调制状态改变的数值,单位是 1/s,称为波特(Baud)。调制速率也称为波特率。波特率描述的是码元传输的速率。

2. 比特率的定义

数据传输速率描述在计算机通信中每秒传送的构成代码的比特数,单位是 bps,因此也可以称为比特率。

在网络问题讨论中,一般都使用比特率(bps)来描述网络的传输性能,只有在讨论到物理层通信技术细节时才有可能涉及波特率的概念。

3. 波特率与比特率的换算

比特率 S(单位为 bps)与调制速率 B(单位为 Baud)之间的关系可以表示为:$S=B\cdot\log_2 k$,式中,k 为多相调制的相数。$\log_2 k$ 值表示一次调制状态的变化传输的二进制比特数。

表 2-3 给出了八相调相绝对调相的相位数值与所表示的 3 位二进制数的对应关系。例如,相位值为 0°表示二进制数 000,45°表示二进制数 001。

表 2-3 八相调相的相位变化值

比特位	相对相位偏移值	比特位	相对相位偏移值
000	0°	100	180°
001	45°	101	225°
010	90°	110	270°
011	135°	111	315°

如果调制速率为 2400Baud，那么多相调制的波特率与比特率的关系如表 2-4 所示。

表 2-4 波特率与比特率的关系

调制速率/Baud	多相调制的相数	$\log_2 k$ 值	数据传输速率/bps
2400	二相调制($k=2$)	1	2400
2400	四相调制($k=4$)	2	4800
2400	八相调制($k=8$)	3	7200
2400	十六相调制($k=16$)	4	9600

表中所示，在 QPSK 调制方法中，当调制速率为 2400(Baud)，多相调制的相数 $k=8$ 时，$\log_2 8=3$ 表示调制解调器的相位状态每变化一次，传输 3 位的二进制数，因此数据传输速率应该为 7200bps。

实际应用中人们经常将不同的调制方法组合起来，以提高频带通信中的数据传输速率。例如，将 ASK 与 PSK 方法相结合，形成正交振幅调制(Quadrature Amplitude Shift Keying，QASK)方法。例如，QAM-64、QAM-128 的编码方法。如果在 2400 波特的线路上使用 QAM-64 编码方法，数据传输速率可以达到：$2400\times\log_2 64=2400\times6=14.49$(kbps)。

问题 2-17：如何理解“带宽”的概念？

计算机与计算机之间通过通信信道传输数据的结构如图 2-10 所示。通信信道应该理解为通信线路与通信设备的集合。所谓通信信道的带宽是指通信线路与通信设备总体能够为传输数据信号提供的带宽。严格地说，通信信道对其所传输的数据信号的影响取决于它的幅频特性。通信信道的幅频特性也是用频率响应曲线来描述的。

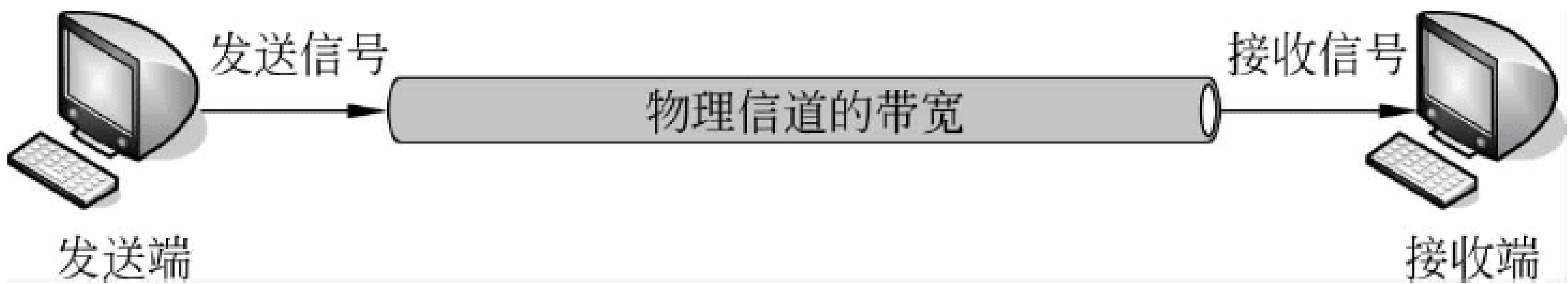


图 2-10 计算机通信系统结构示意图

在现代网络技术中，人们常常使用“带宽(Broad)”这个术语。为了理解“带宽”对网络与通信的重要性，先要了解以下三个问题：什么是通信信道的频率特性？什么是通信信道的带宽？通信信道带宽对基带传输有什么影响？

我们可以回忆一下在电子线路课程中，曾经做过的用于测量一个单级放大器频率特性的实验(如图 2-11 所示)。这个实验的目的是定量测定放大器对不同频率的输入信号的响应特性。在实验中，可以通过调节信号发生器，保持输入的正(余)弦信号的频率可变、幅度不变。通过连接在输出端的电压表读取输出正(余)弦信号的幅度，可以取得对应于不同信

号频率的输出信号电压值 $V_o(f)$,$V_o(f)$曲线叫作该放大器的幅频曲线。通常,人们将放大器输出功率从最大值下降一半所对应的频率范围,定义为放大器的通频带宽,即带宽。上述测量方法与带宽的定义,同样也适合于通信信道。然而,在计算机网络技术的讨论中,“带宽”的概念被扩展了,几乎成了“数据传输速率”的代名词。这个问题在后面要进行专题讨论。

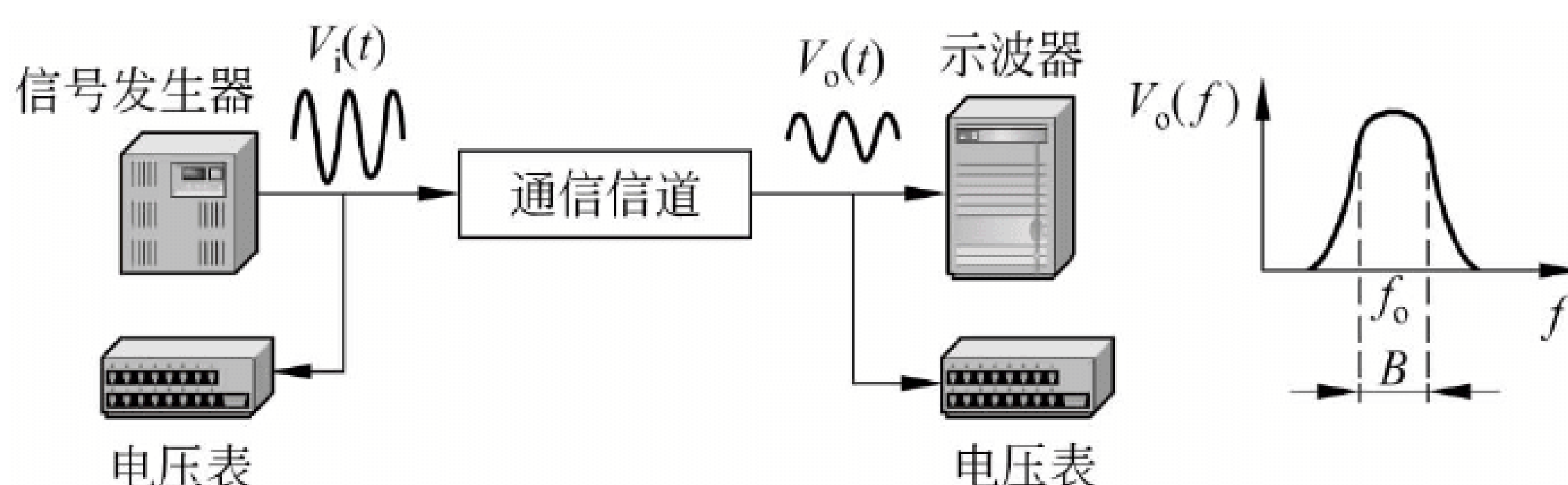


图 2-11　用于测量一个单级放大器频率特性的实验

问题 2-18:如何认识基带信号的频谱特性?

在基带传输的教学中有三个基本的问题需要搞清楚。这三个问题是:基带信号的频谱特性、通信信道的频率特性与通信信道的频带对数据通信速率的影响。

计算机中的数据是用离散的二进制数0、1表示。计算机通信过程中传输的数据流是随机的二进制比特序列。设计适用于计算机通信的数据通信系统,首先要讨论计算机数据的特征及对系统的要求,分析工具是傅里叶级数。

1. 观察与分析电信号的时域方法与频域方法

从电子学的角度,观察与分析电信号的基本方法有两种:时域方法与频域方法。

将时间 t 作为自变量,将电压 u 作为应变量的函数 $u(t)$,去研究电压 u 随着时间 t 变化规律的方法叫作时域方法。如果将角频率 ω 作为自变量,将振幅作为应变量的函数 $E(\omega)$,去研究振幅 E 与角频率 ω 关系的方法叫作频域方法。因此,如果一个电信号用 $u(t)=A\sin(\omega_1 t+\varphi_0)$表示,那么从时域方法看,它是一个振幅为 A、角频率为 ω_1、初位相为 φ_0 的正弦波。如果从频域方法去看,它是一个具有单一角频率为 ω_1 的信号。按照时域方法去表示,数字数据信号是一个随时间跳变的矩形脉冲信号。而从频域方法去研究,它是一个包括直流分量和基波、高次谐波等多个频谱分量的复杂信号。

2. 傅里叶分析的基本方法

频域方法通过傅里叶分析,以频率 f(或角频率 ω,$\omega=2\pi f$)为自变量,观察电信号的频谱组成、振幅与相位,分析电信号通过某一通信信道后频谱分量的变化,最终给出电信号波形的变化及引起传输失真的情况。在数据通信中,表示计算机二进制的比特序列的数字数据信号是典型的矩形脉冲信号。傅里叶分析是通信与电子学的基本分析方法。傅里叶分析指出,任何一个基频为 f 的周期函数 $E(t)$ 都可以表示为无数个正弦和余弦函数之和,它可能写为:

$$E(t)=\frac{a_0}{2}+\sum_{n=1}^{\infty}[a_n\cos(2\pi nft)+b_n\sin(2\pi nft)] \tag{2-1}$$

其中:

$$a_0=\int_0^T E(t)\mathrm{d}t$$

$$a_n = \frac{2}{T}\int_0^T E(t) \cdot \cos(2\pi nft)\mathrm{d}t$$

$$b_n = \frac{2}{T}\int_0^T E(t) \cdot \sin(2\pi nft)\mathrm{d}t$$

式中，a_0是一个常数，f 为基频，周期 $T=1/f$，a_n、b_n分别为 n 次谐波的正弦和余弦信号的振幅值。

傅里叶分析表明：一个周期性函数可以是无数个正弦与余弦函数之和，反之用无数个正弦、余弦函数之和也可以构成一个周期性函数。

3. 周期性矩形脉冲信号的频谱分析

如果周期性矩形脉冲信号的脉冲幅度为 A，带宽为 τ，重复周期为 T，其表达式为：

$$E(t) = \begin{cases} A & \left(nT - \frac{\tau}{2} \leqslant t \leqslant nT + \frac{\tau}{2}\right) \\ 0 & \left((n-1)T + \frac{\tau}{2} < t < nT - \frac{\tau}{2}\right) \end{cases} \tag{2-2}$$

根据式(2-1)可以计算出：

$$a_0 = \frac{2}{T}\int_{-T/2}^{T/2} A \cdot \mathrm{d}t = \frac{2A\tau}{T}$$

$$a_n = \frac{2}{T}\int_{-T/2}^{T/2} A \cdot \cos(2\pi nft) \cdot \mathrm{d}t = \frac{2A\tau}{T} \cdot \frac{\sin(n\pi\tau/T)}{(n\pi\tau/T)}$$

$$b_n = \frac{2}{T}\int_{-T/2}^{T/2} A \cdot \sin(2\pi nft) \cdot \mathrm{d}t = 0$$

那么，周期性矩形脉冲信号可以写为：

$$E(t) = \frac{A\tau}{T} + \sum_{n=1}^{\infty} \frac{2A\tau}{T} \cdot \frac{\sin(\pi nft/T)}{(n\pi\tau/T)} \cdot \cos(2\pi nft) \tag{2-3}$$

从式(2-3)中可以看出，周期性矩形脉冲信号从频域角度观察是一种线谱。它是由直流分量、基波 f 与高次谐波组成。

4. 示例

一个有限宽度的周期数据信号可以被看成一个周期函数。“有限宽度”的概念是 $\tau<<A$。图 2-12 给出了一个周期性发送 1000010000100000…数字数据信号的波形。

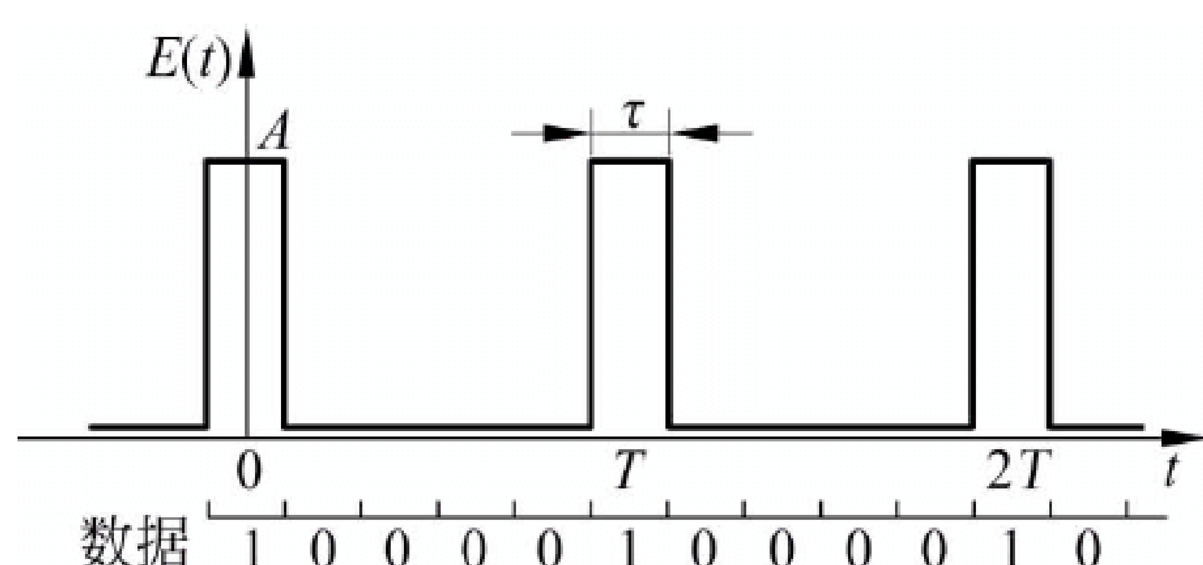


图 2-12　周期性数据脉冲信号波形

如果发送的速率是每秒钟 2000b，那么：

(1) 每比特发送时间　　$\tau=1/2000(\mathrm{s})$

(2) 周期　　$T=5\tau=1/400(\mathrm{s})$

(3) 重复频率　　$f_0=400(\mathrm{Hz})$

(4) 比值 $\tau/T=1/5$

将以上条件代入式(2-3),可以求得:

(1) 直流分量幅值 $a_0=A/5=0.2A$

(2) 基波分量幅值 $a_1=(2A/\pi)\sin(\pi/5)\approx 0.375A$

(3) 二次谐波分量幅值 $a_2=(A/\pi)\sin(2\pi/5)\approx 0.303A$

(4) 三次谐波分量幅值 $a_3=(2A/3\pi)\sin(3\pi/5)\approx 0.202A$

(5) 四次谐波分量幅值 $a_4=(A/2\pi)\sin(4\pi/5)\approx 0.094A$

(6) 五次谐波分量幅值 $a_5=(A/2\pi)\sin(\pi)=0$

(7) 六次谐波分量幅值 $a_6=(A/3\pi)\sin(6\pi/5)\approx -0.060A$

…

(8) 十次谐波分量幅值 $a_{10}=(A/5\pi)\cdot\sin 2\pi=0$

…

那么周期性发送的10000数据脉冲信号可以写为:

$$
\begin{aligned}
E(t)=&0.2A && \text{直流分量}\\
&+0.375A\cos(\omega_0 t) && \text{基波分量}\\
&+0.303A\cos(2\omega_0 t) && \text{二次谐波分量}\\
&+0.202A\cos(3\omega_0 t) && \text{三次谐波分量}\\
&+0.094A\cos(4\omega_0 t) && \text{四次谐波分量}\\
&+0 && \text{五次谐波分量}\\
&+0.06A\cos(6\omega_0 t+\pi) && \text{六次谐波分量}\\
&\cdots\\
&+0 && \text{十次谐波分量}\\
&\cdots
\end{aligned}
$$

式中,$\omega_0=2\pi f_0$,$f_0=400\text{Hz}$。图2-13给出了该信号的频谱分布图。

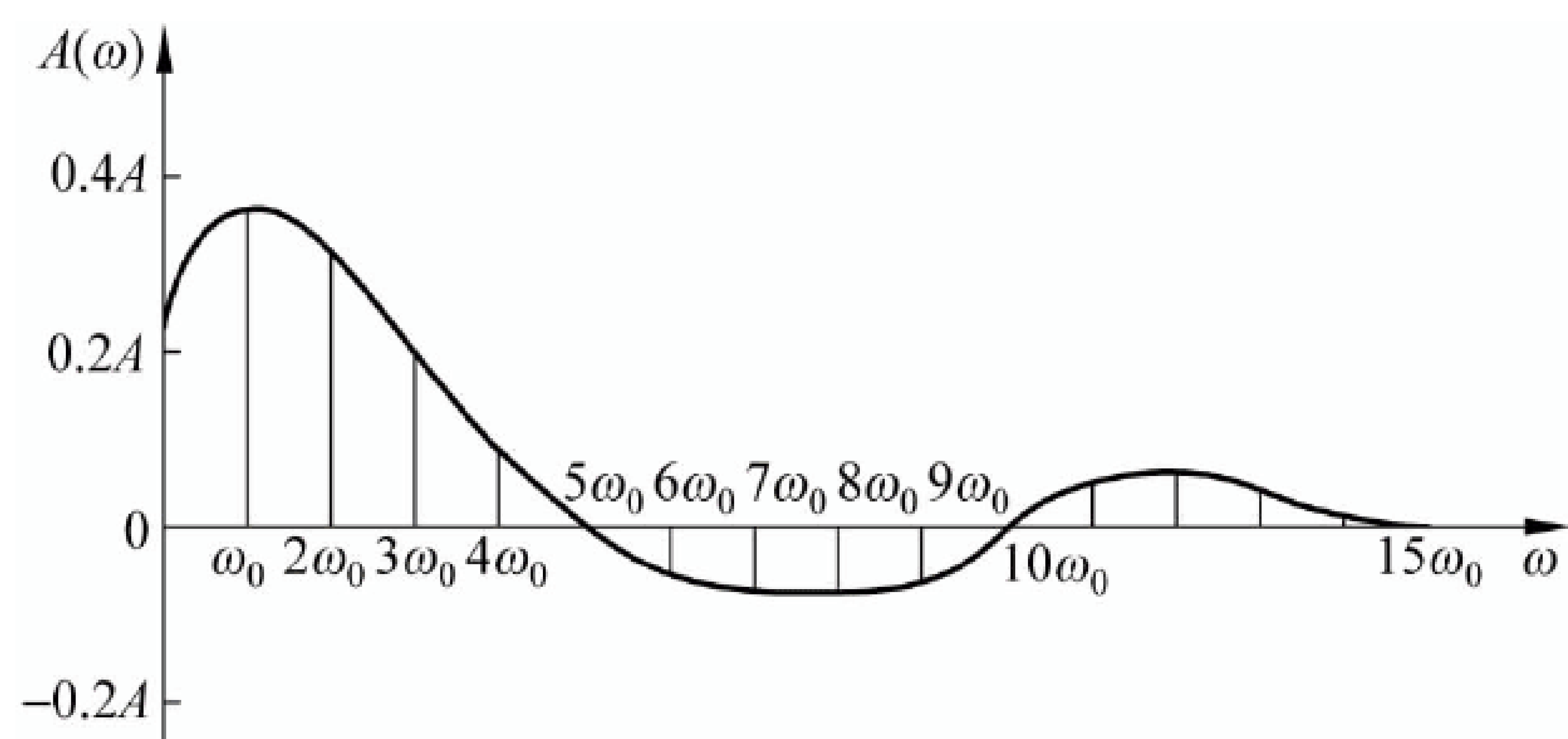


图2-13 周期性数字脉冲信号频谱

5. 数字信号的频谱特点

从以上讨论中可以看出,周期性数字信号的频谱具有以下特点。

(1) 它是由直流分量、基波分量以及高次谐波分量组成,其频谱为线谱;

(2) 频谱分量幅度的包络呈 $\sin(x)/x$ 的形式,$x=n\pi\tau/T$,当 $x=\pi,2\pi,3\pi,\cdots$ 时,包络出现零点;

(3) 包络的第一个零点近似等于传输该信号所需要的通信信道带宽 B,$B=nf_0$,

$n=T/\tau$。

分析说明：周期性数字信号的频谱占有一定的宽度，且与脉冲波形有关。脉冲越窄的信号频谱的宽度越宽，传输时所需要通信信道的带宽就越宽。为了使数据信号通过通信信道的失真减小，就要求根据数字信号的频谱与传输速率来选择通信信道的带宽。

问题 2-19：信道带宽对基带信号传输有什么样的影响？

为研究通信信道带宽对数据信号传输的影响，我们以连续发送一个 8 比特组成的字符 B 为例说明。字符 B 的二进制 ASCII 码为 01100010，如果重复发送该字符，可以将该数据信号视为一个周期函数。图 2-14(a)给出作为发送数据 B 的脉冲信号波形与频谱。图 2-14(b)表示：如果通信信道带宽较窄，只允许发送信号的直流分量与基波分量通过时，用傅里叶积分的方法，得到通过通信信道后，由直流分量与基波分量合成的接收信号波形，结果说明：接收信号波形失真很大，接收端无法识别信号的正确编码。图 2-14(c)表示：如果通信信道带宽允许发送信号的直流分量、基波分量与二次谐波分量通过时，接收信号的波形仍然失真很大。图 2-14(d)表示：如果通信信道带宽允许发送信号的直流分量、基波分量、二次与四次谐波分量通过时，接收信号波形已开始接近发送信号波形。图 2-14(e)表示：如果通信信道带宽较宽，允许发送信号的直流分量、基波分量、二次至八次谐波通过时，接收信号波形已比较接近发送信号波形，接收方经波形整型后可以正确识别 8 位二进制值。

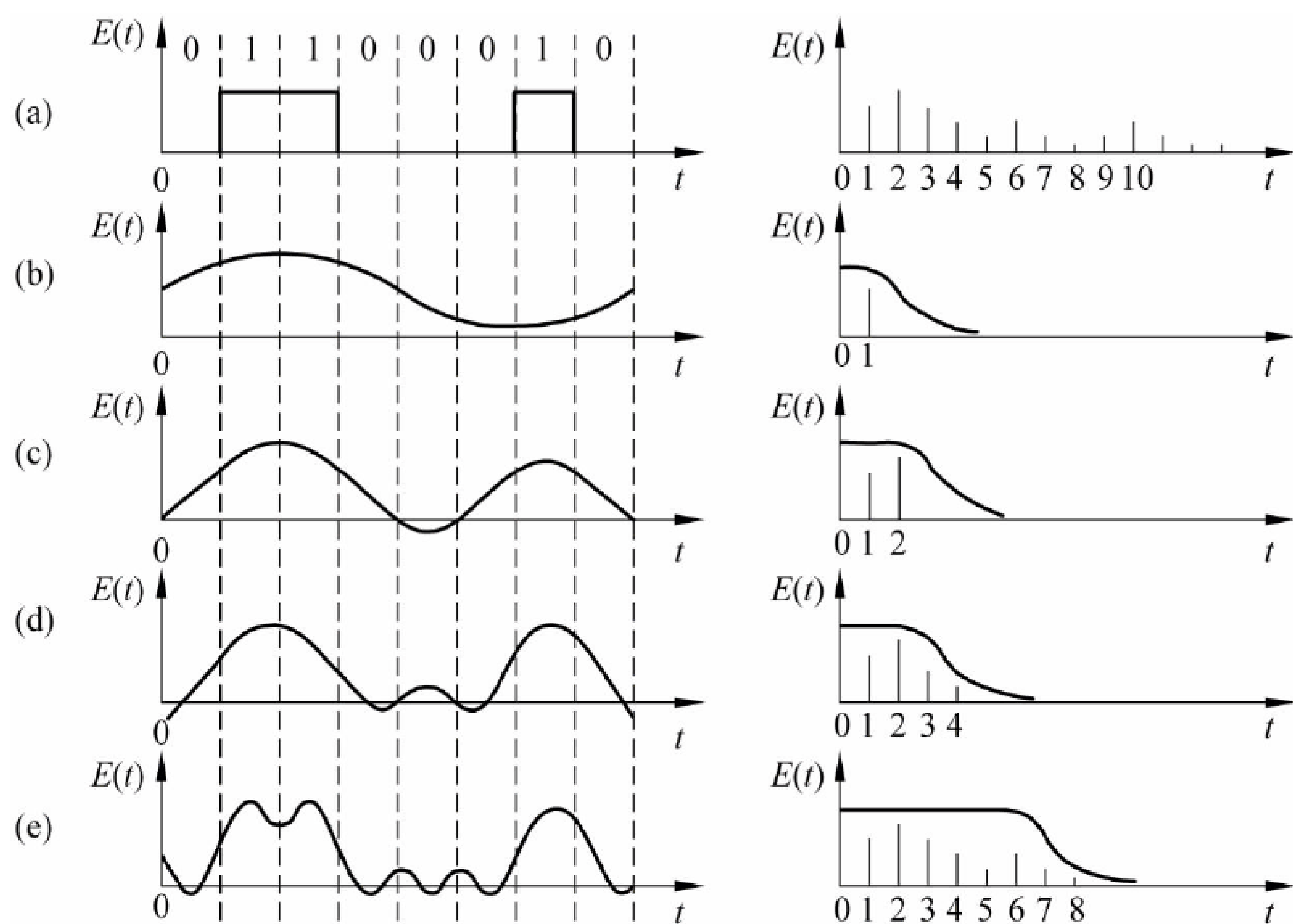


图 2-14　通信信道带宽对数字数据信号接收波形的影响

问题 2-20：为什么不同教科书的曼彻斯特编码波形可能是不同的？

解释这个问题需要注意以下几个问题。

1. 曼彻斯特编码(Manchester Coding)规则

曼彻斯特编码是目前应用最广泛的编码方法之一。曼彻斯特编码规则如下。

(1) 每比特的周期 T 分为前 $T/2$ 与后 $T/2$。

(2) 前 $T/2$ 传送该比特的反码。

(3) 后 $T/2$ 传送该比特的原码。

根据曼彻斯特编码规则,如图2-15所示:$b_0=0$,它的前$T/2$取0的反码。0用低电平表示,其反码为高电平;后$T/2$取0的原码(低电平)。$b_1=1$,前$T/2$取1的反码低电平;后$T/2$取1的原码(高电平)。$b_2=0$,b_2的前$T/2$为高电平,后$T/2$为低电平。$b_3=0$,b_3的前$T/2$为高电平,后$T/2$为低电平。按这个规律可以画出图中曼彻斯特编码信号的波形图。

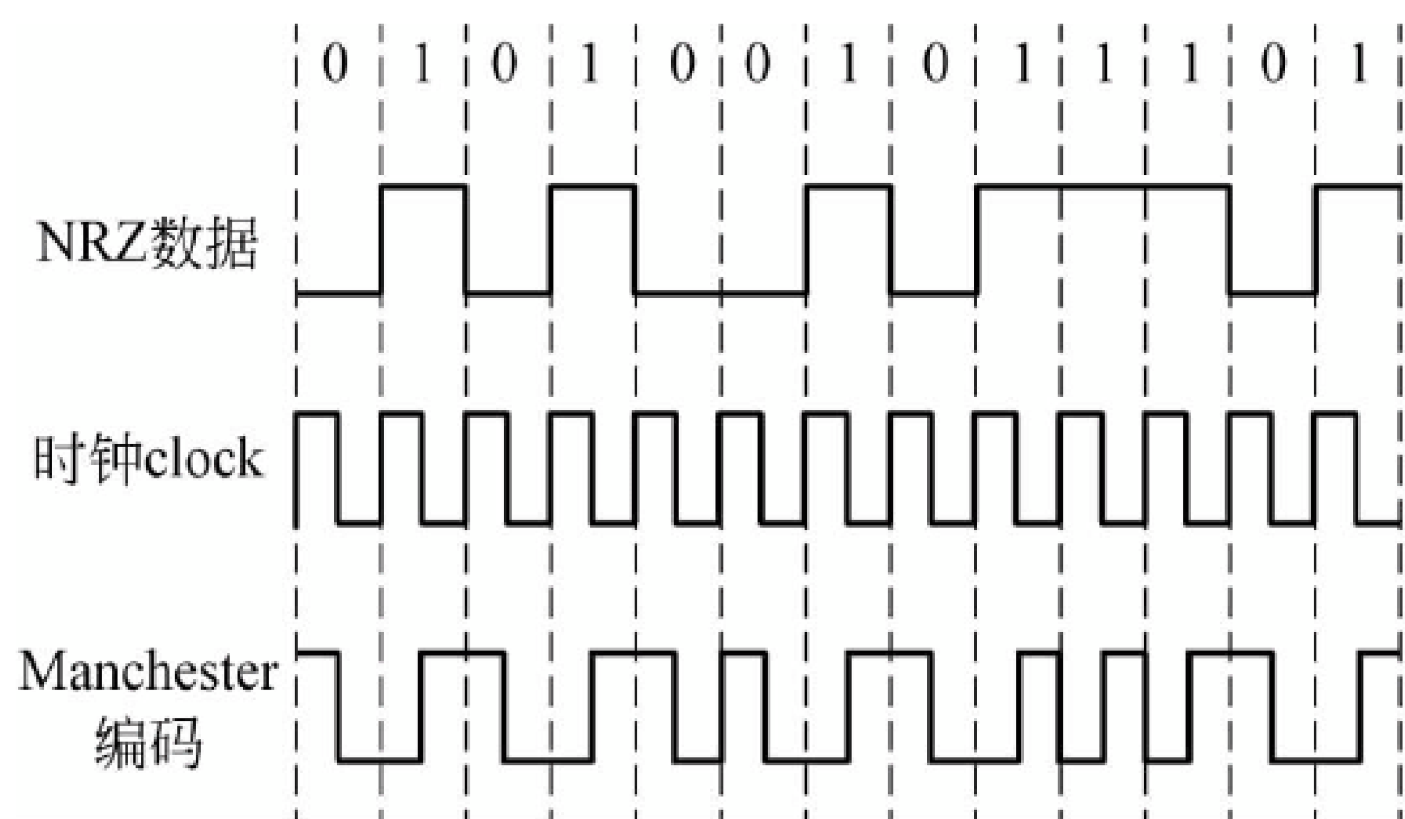


图2-15　曼彻斯特编码信号波形图

2. IEEE 802.3标准对曼彻斯特编码的说明

IEEE 802.3标准规定曼彻斯特编码的规则是:数据与时钟进行“异或”运算,因此就造成了每比特前$T/2$取该比特的反码,后$T/2$传送该比特的原码。不同的教科书在曼彻斯特编码波形的表述中存在着两种方法,差别是在第一个码元的前$T/2$是取反码还是原码上。有些教科书采用了第一个码元的前$T/2$取原码的方法,因此会造成曼彻斯特编码信号波形上的差异。

问题2-21:为什么要研究脉冲编码调制PCM技术?

1. 脉冲编码调制PCM的作用

由于数字信号传输失真小、误码率低、数据传输速率高,因此在网络中除计算机直接产生的数字外,语音、图像信息的数字化已成为发展的必然趋势。脉冲编码调制PCM是实现模拟数据数字化的主要方法。

PCM技术的典型应用是语音数字化。语音可以用模拟信号的形式通过电话线路传输,但是在网络中将语音与计算机产生的数字、文字、图形与图像同时传输,就必须首先将语音信号数字化。在发送端通过PCM编码器将语音信号变换为数字化语音数据,通过通信信道传送到接收端,接收端再通过PCM解码器将它还原成语音信号。数字化语音数据的传输速率高、失真小,可以存储在计算机中,并且进行必要的处理。因此,在网络与通信中,首先要利用PCM技术将语音数字化。图2-16是模拟语音信号通过脉冲编码调制技术形成数字语音信号的过程示意图。

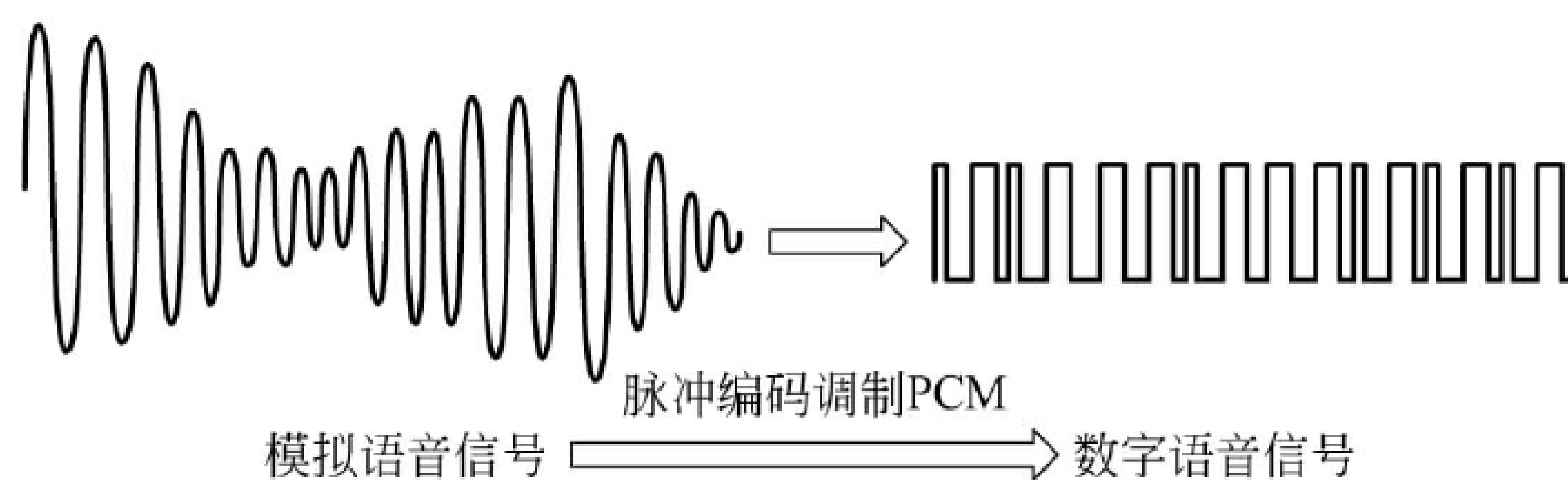


图2-16　脉冲编码调制示意图

2. 脉冲编码调制基本工作原理

脉冲编码调制方法的基本工作原理包括三部分内容：采样、量化与编码。

1）采样

模拟信号是电平连续变化的信号，采样是隔一定的时间间隔，将模拟信号的电平幅度值取出作为样本，让其表示原来的信号。采样频率 f 应为：

$$f \geqslant 2B \quad \text{或} \quad f = 1/T \geqslant 2f_{\max}$$

式中，B 为通信信道带宽，T 为采样周期，$f_{\max}$为信道允许通过的信号最高频率。

理论研究结果表明，如果以大于或等于通信信道带宽二倍的速率定时对信号进行采样，其样本可以包含足以重构原模拟信号的所有信息。

2）量化

量化是将采样样本幅度按量化级决定取值的过程。经过量化后的样本幅度为离散的量化级值，已不是连续值。量化之前要规定将信号分为若干量化级，例如，可以分为 8 级或 16 级，以及更多的量化级，这要根据精度要求决定。同时，要规定好每一级对应的幅度范围，然后将采样所得样本幅值与上述量化级幅值比较。

3）编码

编码是用相应位数的二进制代码表示量化后的采样样本的量级。如果有 k 个量化级，则二进制的位数为 $\log_2 k$。

当 PCM 用于数字化语音系统时，它将声音分为 128 个量化级，每个量化级采用 8 位二进制编码表示。由于采样速率为 8000 样本/秒，因此数据传输速率应达到 8×8000＝64(kbps)。

PCM 还可以用于计算机中的图形、图像数字化与传输处理中。PCM 采用二进制编码的缺点是使用的二进制位数较多，而编码效率较低。

问题 2-22：奈奎斯特准则是如何推导出来的？

任何通信信道都不是理想的，也就是说，信道带宽总是有限的。由于信道带宽的限制、信道干扰的存在，信道的数据传输速率总会有一个上限。早在 1924 年奈奎斯特(Nyquist)就推导出具有理想低通矩形特性的信道，在无噪声情况下的最高速率与带宽关系的公式，这就是奈奎斯特准则。要理解奈奎斯特准则的结论，必须注意以下几个基本问题。

1. 奈奎斯特准则的推导条件

1）理想通信信道的频率响应曲线

理想通信信道的频率响应曲线如图 2-17 所示。

$$H(\omega) = \begin{cases} e^{-j\omega\tau} & (-\omega_1 \leqslant \omega \leqslant +\omega_1) \\ 0 & (+\omega_1 < \omega, \omega < -\omega_1) \end{cases}$$

2）讨论单个矩形脉冲通过理想通信信道的方法

图 2-18 给出了单个矩形脉冲通过理想通信信道的示意图。

2. 推导过程

利用傅里叶分析可以求出通过理想通信信道的接收信号：

$$H(t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} H(\omega) \cdot e^{j\omega t} dt = \frac{1}{2\pi}\int_{-\omega_1}^{+\omega_1} e^{-j\omega\tau} \cdot e^{j\omega t} dt$$

$$= \frac{1}{\pi}\int_0^{\omega_1}\cos(t-\tau)\mathrm{d}\omega = \frac{\omega_1}{\pi}\cdot\frac{\sin\omega_1(t-\tau)}{\omega_1(t-\tau)}$$

$$= \frac{\omega_1}{\pi}\cdot\frac{\sin\omega_1 x}{x}$$

式中,$x=\omega_1(t-\tau)$。$H(t)$波形如图2-19所示。

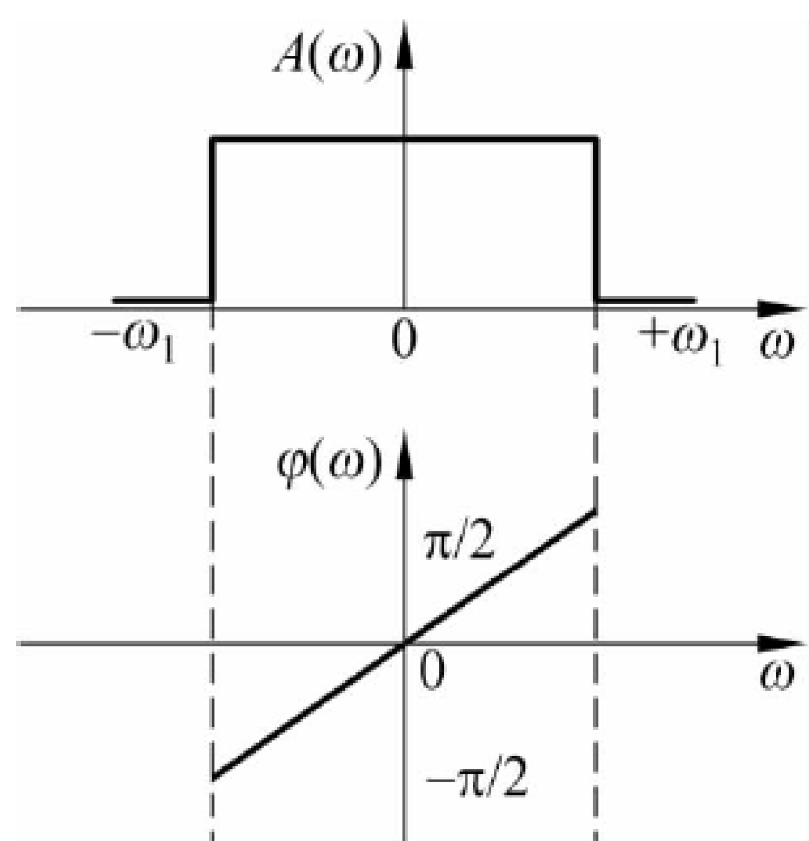


图2-17　理想通信信道的频率响应曲线

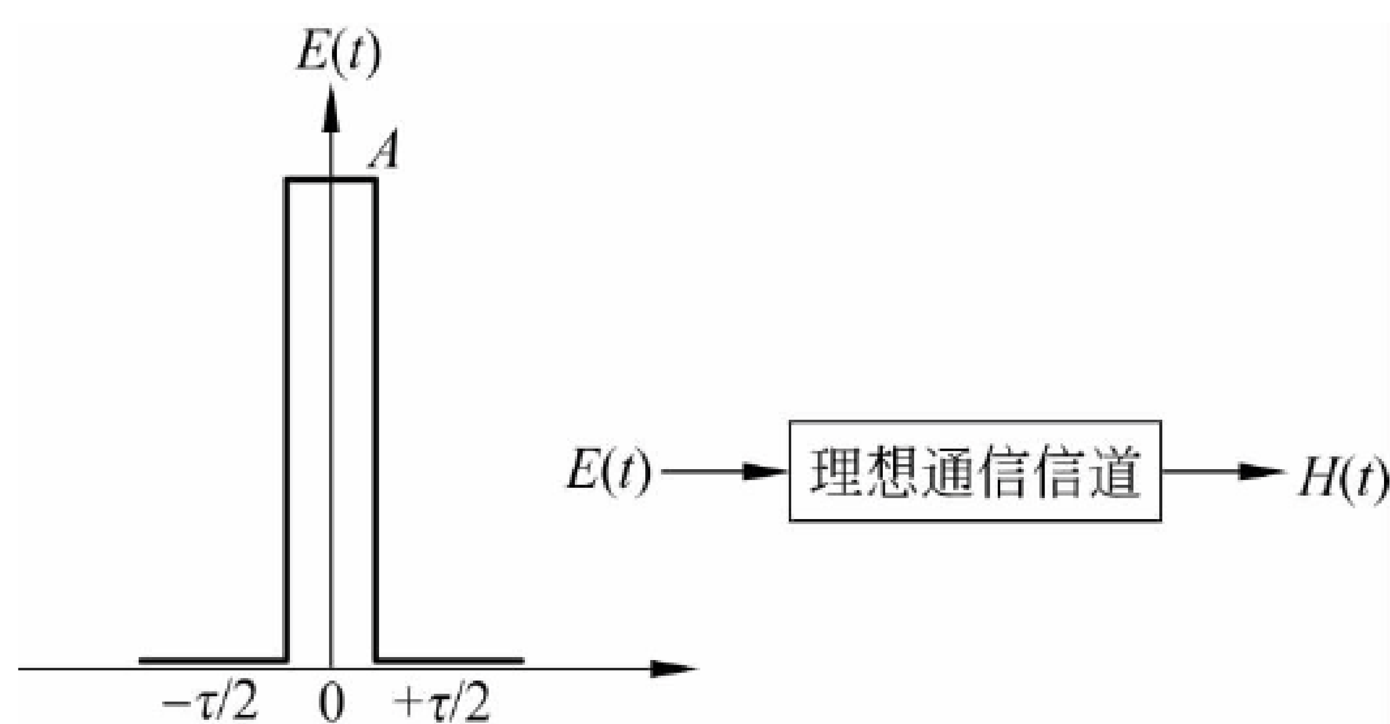


图2-18　单个矩形脉冲通过理想通信信道示意图

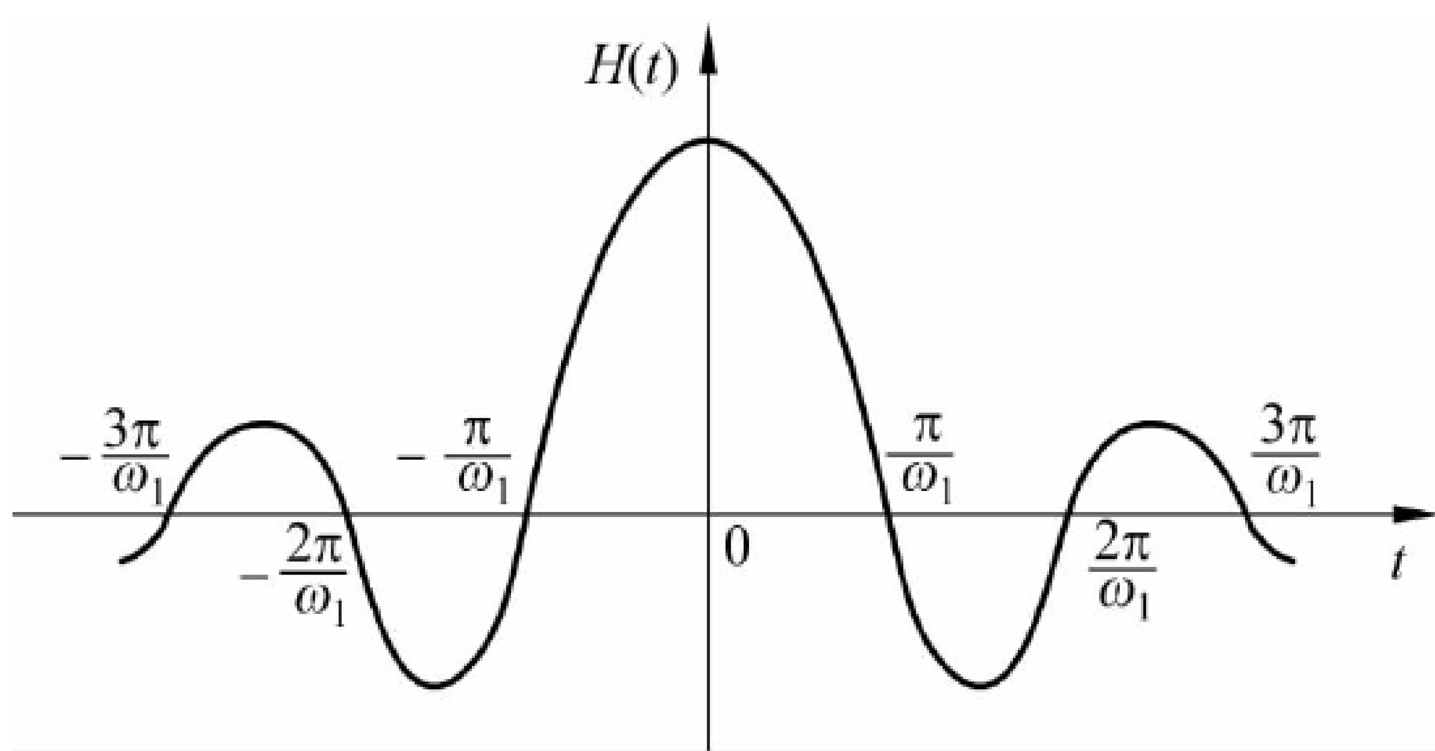


图2-19　$H(t)$波形图

窄矩形脉冲信号通过理想幅频与相频的通信信道后,接收端出现的信号波形呈 $\sin x/x$ 形。每隔 π/ω_1时间间隔出现零点。信号能量主要集中在$+\pi/\omega_1$与$-\pi/\omega_1$之间。

3. 结论

奈奎斯特准则指出:如果间隔为 $\pi/\omega(\omega=2\pi f)$,通过理想通信信道传输窄脉冲信号,则前后码元之间不产生相互串扰(如图2-20所示)。

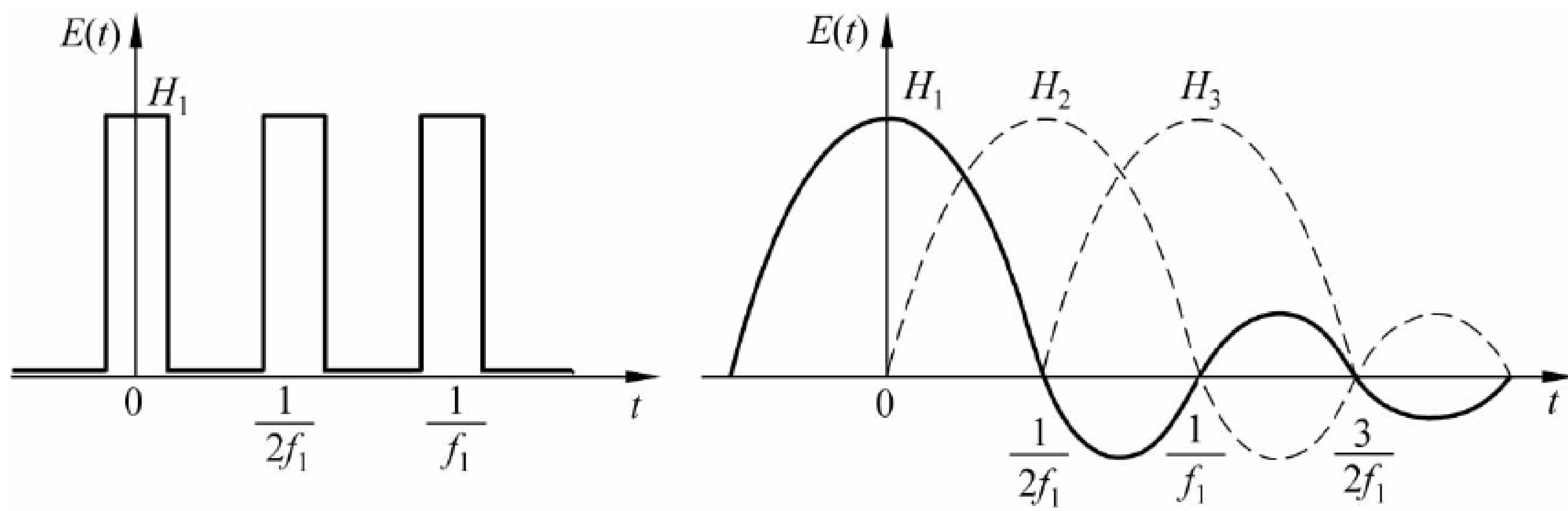


图2-20　窄脉冲信号通过理想通信信道

因此,对于二进制数据信号的最大数据传输速率 R_{max}与通信信道带宽 B(单位Hz)的关系可以写为

$$R_{max} = 2B(\text{bps})$$

对于二进制数据，若信道带宽 $B=f=3000\text{Hz}$，则最大数据传输速率为 6000bps。

问题 2-23：为什么在计算机网络的讨论中可以用带宽来取代速率？

奈奎斯特定理描述了有限带宽、无噪声信道的最大数据传输速率与信道带宽的关系。香农定理则描述了有限带宽、有随机热噪声信道的最大传输速率与信道带宽、信号噪声功率比之间的关系。

香农定理指出：在有随机热噪声的信道上传输数据信号时，数据传输速率 R_{max} 与信道带宽 B，信号与噪声功率比 S/N 关系为

$$R_{max} = B \cdot \log_2(1+S/N)$$

式中，R_{max} 单位为 bps，带宽 B 单位为 Hz，信号与噪声功率比（简称信噪比），通常以 db（分贝）数表示。若 $S/N=30(\text{db})$，那么信噪比根据公式：

$$S/N(\text{db}) = 10 \cdot \lg(S/N)$$

可得 $S/N=1000$。若带宽 $B=3000\text{Hz}$，则 $R_{max}\approx 30(\text{kbps})$。香农定律给出了一个有限带宽、有热噪声信道的最大数据传输速率的极限值。它表示对于带宽只有 3000Hz 的通信信道，信噪比在 30dB 时，无论数据采用二进制或更多的离散电平值表示，都不能用超过 30kbps 的速率传输数据。

从奈奎斯特定理和香农定理的讨论中，我们已经看到通信信道的最大传输速率与信道带宽之间存在着明确的关系，人们可以用“带宽”去取代“速率”。例如，人们常把网络的“高数据传输速率”用网络的“高带宽”去表述。因此，“带宽”与“速率”在网络技术的讨论中几乎成了同义词。

问题 2-24：为什么 1kbps≠1024bps？

曾经有同学提出：1Kb=1024b，为什么 1kbps≠1024bps 呢？回答这个问题很简单：这个区别是由计算机学科与通信学科所采用的二进制与十进制引起的。

在计算二进制的数据长度时，1Kb=1024b；但在计算通信速率时使用的是十进制，1kbps=1000bps≠1024bps；同样，40.98×10^6 bps=40.98Mbps≠40.00Mbps。

同时还需要注意以下两个问题。

（1）在计算机网络的速率单位的英文标识中，kbps 中的“k”是用小写的英文字母，而 Mbps 与 Gbps 中的“M”与“G”都用大写的英文字母。

（2）数据传输速率的单位“比特/秒”的英文表示有两种：“bps”与“b/s”。在本教材中，作者采用国外经典教材和重要标准中的表示方法，即“bps”。

问题 2-25：多路复用中“帧”与数据链路层中“帧”的区别是什么？

回答这个问题需要注意以下两点。

1. 多路复用中的“帧”

多路复用是属于物理层的问题。我们可以用图 2-21 中的时分多路复用 T1 载波帧结构来说明这个问题。T1 载波帧是针对脉冲编码调制 PCM 时分多路复用而设计，用于在接收端将复用的 24 路信号分离开。

2. 数据链路层中的“帧”

数据链路层中的“帧”是用来实现数据链路层协议的，它用来检查、发现传输的 IP 分组

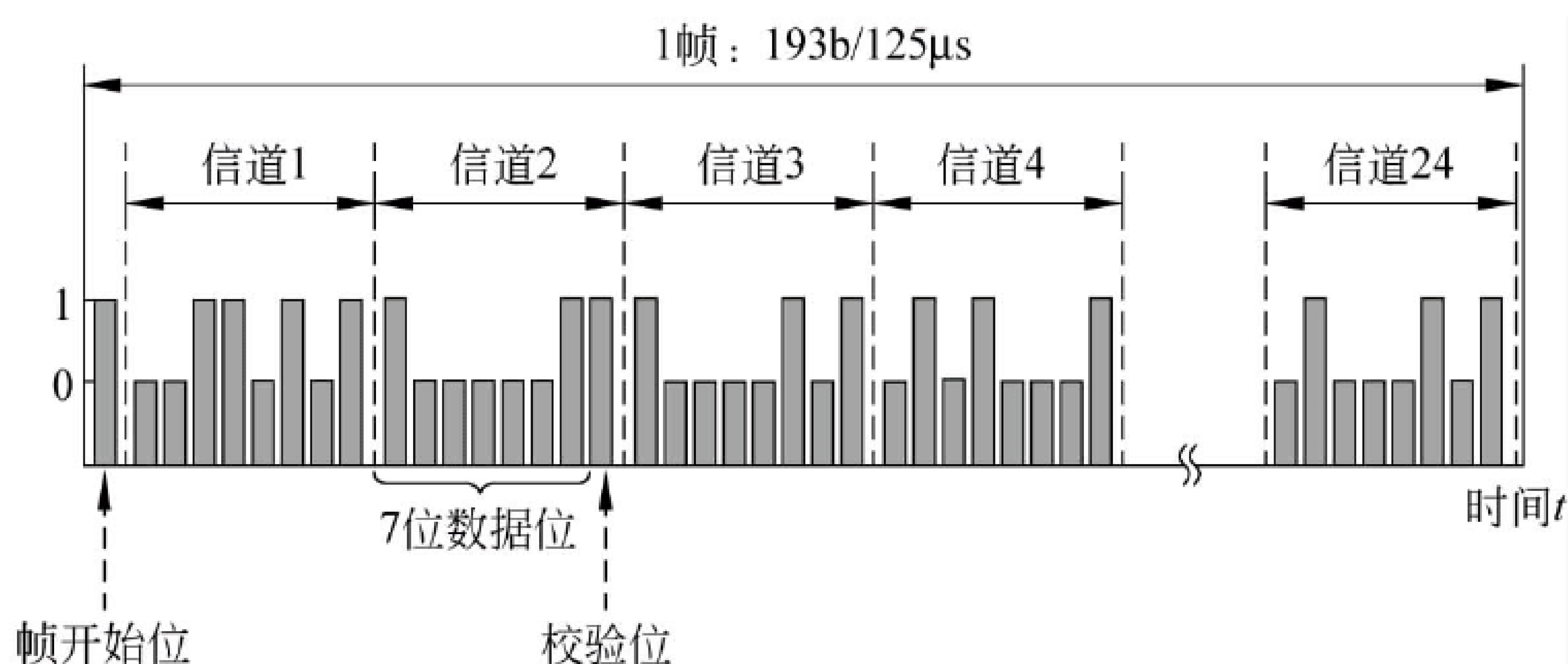


图 2-21　时分多路复用 T1 载波帧结构

的数据是否在物理层传输过程中出现错误。因此数据链路层中的帧需要加上帧头与帧尾。因此，两种数据结构都叫“帧”，但是设置的目的不同、作用不同、结构不同。

问题 2-26：为什么会出现 T1、E1 等多种载波速率体系？

这是由历史原因造成的。由于在研究时分多路复用技术的初期，不同地区分别在研究各自的技术与标准，因此就形成在北美的 T1 载波(T1 Carrier)速率体系、欧洲的 E1 载波(E1 Carrier)速率体系，以及日本的速率标准体系。

北美的 T1 载波基本速率是将 24 路音频信道复用在一条通信线路，速率为 1.544Mbps。欧洲的 E1 载波包括 30 路语音信道和两路传送控制信道；速率为 2.048 Mbps。日本在 T1 载波速率的基础上形成了自己的速率体系。这样才出现了 SONET/SDH 速率体系的研究与标准的制定。

问题 2-27：SONET 是在什么样的背景下发展起来的？

各国的数字传输系统都是采用脉冲编码调制 PCM 与多路复用技术。随着电信网的发展和用户对需求的不断提高，传统的数字传输系统所能提供的带宽，已经远不能适应视频与多媒体传输的要求。基于光纤的高速率通信应用越来越广泛。如果不对高次群的数据传输速率进行标准化，国际通信网络的互联将会十分困难。早期数字传输系统与设备在运行过程中暴露出的问题主要有以下几点。

1. 数据传输速率不标准

由于历史的原因，多路复用的速率体系中存在着两个互不兼容的标准，即北美和日本的 T1 载波，欧洲的 E1 载波。日本又使用了第三种不兼容的标准。

2. 光设备接口标准不规范

高速率的通信主干网将采用光纤，但是在光设备接口方面没有国际的标准规范。各个厂家使用自己的标准，这就造成不同公司的光设备之间的接口困难，增加了网络的复杂性和运营成本。

3. 复用系统中的同步问题

传统的多路复用系统中，为降低设备成本，除了低速率的信号传输中采用同步以外，在其他高速率的复用信号传输中一般采用准同步方式。

要解决这个问题，人们只能从根本上进行改革，研究一种新的技术来取代传统的数字传输速率体系——同步光纤网 SONET。

问题 2-28：如何理解传输网中“同步”的概念？

为了保证数据传输系统的正常工作，必须要求接收端的接收时钟与发送端的发送时钟严格保持一致，否则接收电路就不能正确地判断所接收的二进制比特。同步是保持接收时钟与发送时钟一致性的过程。

1. 同步

如果一组信号为同步信号，意味着信号之间是以绝对相同的速率和相位传输。如果信号之间的相位或速率存在偏差，则这个偏差必须在规定的范围内。在同步网络中，所有时钟都通过基本的参考时钟 PRC 获得，PRC 的精度必须保持在$\pm1\times10^{-11}$内。这样的时钟精度只能通过铯原子钟获得。

2. 准同步

如果一组信号为准同步信号，意味着信号之间的速率和相位必须基本相同。如果信号之间的相位或速率存在偏差，则这个偏差也必须在规定的范围内。在两个互联的网络中，每个网络中的时钟也都通过基本的参考时钟 PRC 获得，但是两个网络的参考时钟 PRC 之间的精度可能存在着偏差，因此这种系统一般称为准同步系统。

3. 异步

如果一组信号为异步信号，意味着各个信号之间的速率和相位偏差要大于准同步信号。如果两个网络的时钟分别是从各自的石英振荡器中获得，则这两组信号就是异步信号。由于异步传输系统的时钟是独立和非同步的，接收时钟与发送时钟的差异就会造成发送速率与接收速率的差异。例如，DS3 的时钟误差为$\pm20\times10^{-6}$，DS3 的速率为 44.734Mbps，时钟误差可能造成的接收速率与发送速率的最大误差为±894.7bps。因此，为了保证接收端能正确识别接收二进制比特流，接收端就必须采用复杂的同步技术。

在传统的多路复用系统中，为了降低设备成本，除了低速率的信号传输采用同步以外，在其他高速率的复用信号传输中一般采用准同步方式。当数据传输速率较低时，收发双方的时钟频率的微小差异不会带来严重的影响。在数据传输速率不断提高时，收发双方的时钟频率的同步问题必须认真加以解决。

问题 2-29：SDH 技术是在什么样的背景下发展起来的？

回答这个问题，需要注意以下几点。

(1) 随着用户对网络的要求在不断变化，现代电信网必须能迅速地为用户提供各种新的通信服务。如果不研究和制定数据传输速率体系标准，国际范围的高速数据传输网络的建设将会非常困难。要想解决这个问题，人们只能从根本上进行改革，研究一种新的技术来取代传统的数字传输速率体系。这种技术就是建立在光纤传输基础上的同步光纤网 SONET。

(2) 同步光纤网的概念由美国贝尔通信研究所首先提出。设计同步光纤网的目的是解决光接口标准规范问题，定义同步传输的线路速率的等级体系，以使不同厂家的产品可以互连，从而能够建立大型的光纤数据传输网络。

(3) 1988 年，ITU-T 接受了 SONET 的概念，并重新命名为同步数字体系 SDH，使它不仅能够适用于光纤，也能够适用于微波和卫星传输，这样就成为通用性技术体制。

(4) ITU-T 推出了一系列有关 SDH 的标准，对 SDH 的速率、复用帧结构、复用设备、

线路系统、光接口、网络管理和信息模型等进行定义,从而确立了作为国际标准的同步数字体系SDH。

问题2-30:SDH传输网具有哪些主要的技术特点?

SDH作为一种传输网技术,主要具有以下几个特点。

(1) STM-1统一了T1与E1载波两大不同的数字速率体系,使数字信号在传输过程中不再需要转换标准,真正实现了数字传输体制上的国际性标准。

(2) SDH网兼容光纤分布式数据接口(FDDI)、分布队列双总线(DQDB),以及ATM信元。

(3) SDH采用同步复用方式,各种不同等级的码流在帧结构负荷内的排列有规律,而净荷与网络是同步的,因此只需利用软件即可使高速信号一次直接分离出低速复用的支路信号,这就降低了复用设备的复杂性。

(4) SDH帧结构的网络管理字节增强了网络管理能力,同时通过将网络管理功能分配到网络组成单元,可以实现分布式传输网络的管理。

(5) 标准的开放型光接口可以在光缆上实现不同公司光接口设备的互连,有效地降低了组网成本。

总结以上的介绍可以看出,SDH传输网最具竞争力的特点是:同步复用、标准光接口与网络管理能力,这些特点决定了SDH网能够成为理想的广域网、城域网的数据传输平台。SONET/SDH标准已成为国际公认的传输网速率体制,它对推动世界电信网络的发展有着重要的作用。

问题2-31:如何理解SONET同步封装净荷的概念?

我们可以用SONET速率标准的STS-1速率51.840Mbps为例来说明这个问题。

1. STS-1帧结构

STS-1帧是一个块状结构,每秒钟发送8000帧。帧的总长度为810B,其中27B用于线路管理开销。用于传输用户数据的部分称为同步封装净荷(Synchronous Payload Envelope,SPE)。STS-1帧中SPE长度为783B。为了表示方便,通常将一个STS-1帧画成9行90列的结构。图2-22给出了STM-1帧结构。

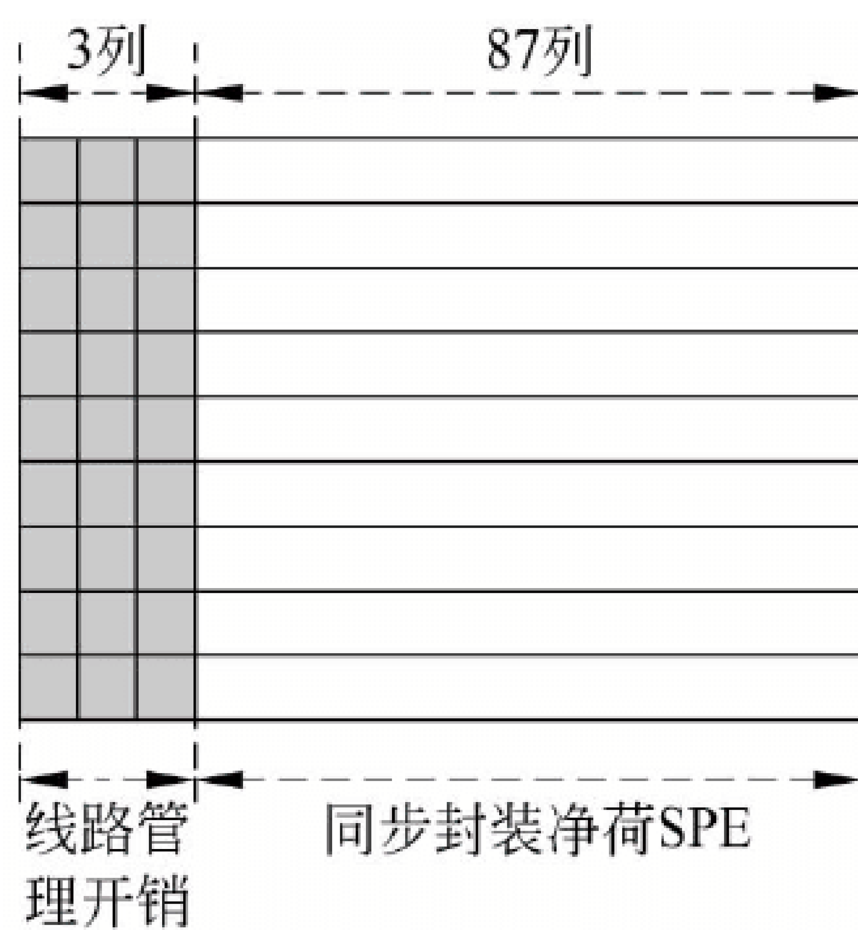


图2-22　STM-1帧结构

2. STM-1的传输参数

根据以上提供的参数可以计算出:

(1) 总的传输速率:8×9×90×8000=51.840(Mbps)

(2) 线路管理消耗带宽:8×3×9×8000=1.728(Mbps)

(3) 同步封装净荷速率:8×9×87×8000=50.112(Mbps)

从以上数据中可以看出,真正能够用于传输用户数据的是同步封装净荷速率,即50.112Mbps。

STM-3及其他帧结构与同步封装净荷速率的分析方法是相同的。

问题2-32:应对不同应用环境需求的接入技术主要有哪些类型?

图2-23给出了不同应用环境的接入技术的类型与相关的标准。从用户接入的角度,可

以分为接入技术与接入方式，接入方式与用户工作环境与需求相关。

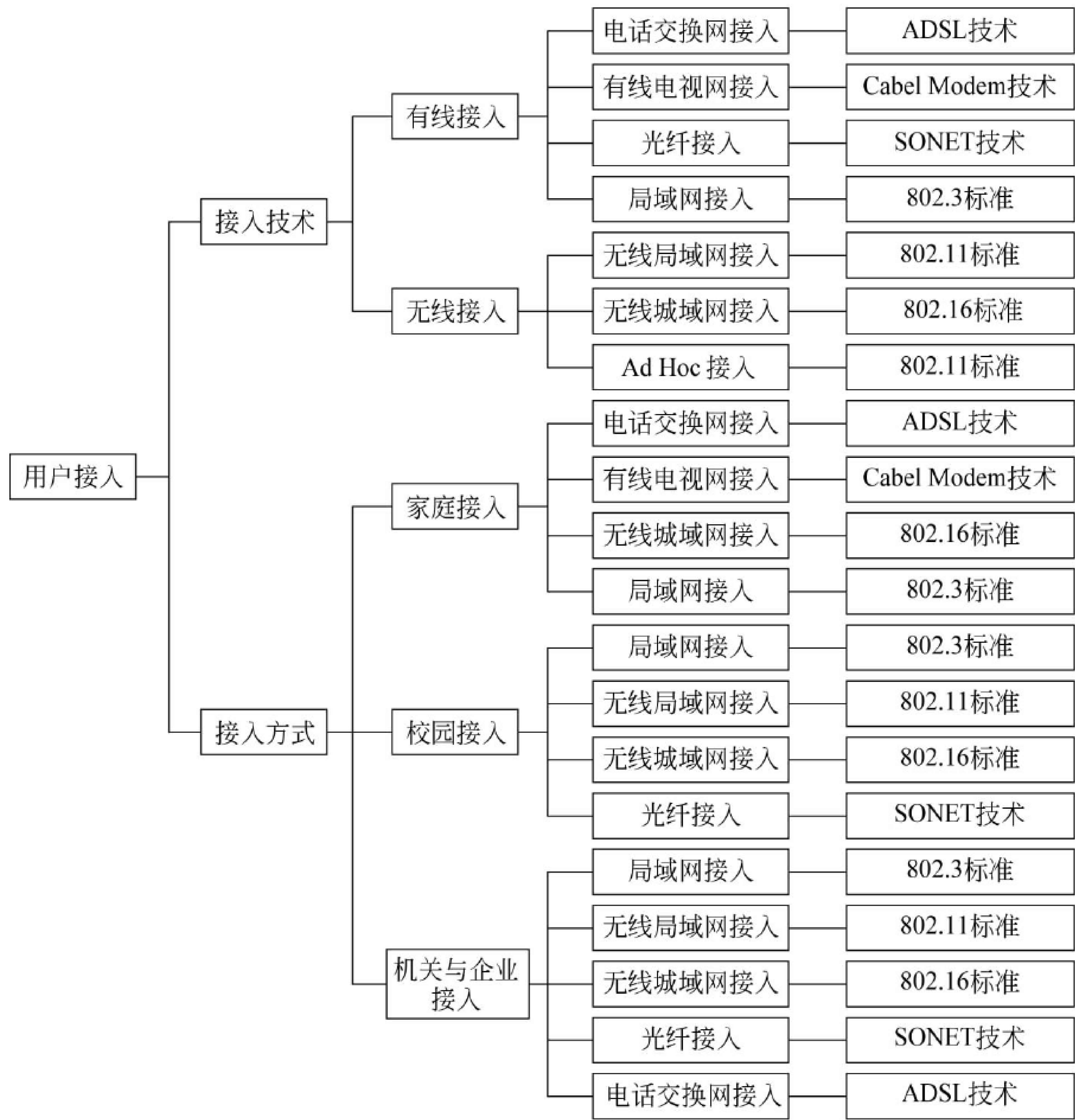


图 2-23　接入技术的类型

接入技术可以分为有线接入与无线接入两类；接入方式涉及用户的环境与需求，它大致可以分为家庭接入、校园接入、机关与企业接入。

从实现技术的角度，目前宽带接入技术主要有以下几种：数字用户线（xDSL）、光纤同轴电缆混合网（HFC）、光纤接入、无线接入与局域网接入。无线接入又可以分为无线局域网接入、无线城域网接入与无线 Ad Hoc 接入。

问题 2-33：如何理解数字用户线 xDSL 接入技术？

认识数字用户线 xDSL 接入技术需要注意以下几个问题。

1．xDSL 的基本概念

大多数电话公司倾向于推动数字用户线（Digital Subscriber Line，xDSL）的应用。数字用户线又叫作数字用户环路。数字用户线是指从用户到本地电话交换中心的一对铜双绞线，本地电话交换中心又叫作中心局。xDSL 是美国贝尔通信研究所于 1989 年为推动视频点播（VOD）业务开发出的基于用户电话铜双绞线的高速传输技术。

电话网是唯一可以在几乎全球范围内向住宅和商业用户提供接入的网络。据估计，全

球电话用户的总数约为7亿。电话通过铜双绞线连接用户家庭与办公室。铜双绞线最初的设计是用于传输模拟话音信号,使用调制解调器 Modem 后,也可以传输数据信号。目前,调制解调器的传输速率可以达到56kbps。由于电话交换网络及调制解调器的限制,进一步提高传输速率是很困难的。在20世纪80年代,ISDN 利用一对双绞线实现了传输速率为144kbps,传输距离为6000m的数据传输。它将144kbps分为两个64kbps的交换D信道和一个16kbps的信令B信道。ISDN 的应用并不是很成功。随着 Internet 的迅速发展,用户对固定结点的宽带接入的需求也日益增加。电信企业的主干网已采用2.5Gbps和10Gbps的超高速光纤,但由于连接用户和交换局的用户线绝大多数仍是电话铜双绞线,以现有的调制技术不能满足用户高速接入的需求。采用 xDSL 技术后,可以在电话铜双绞线上传送高达数 Mbps 速率的数字信号。如果配置了分离音频频带和高频带的分离器,可以同时提供电话和高速数据业务。基于铜双绞线的 xDSL 技术以低成本实现用户线高速传输而异军突起,打破了宽带通信由光纤独揽的局面。

2. xDSL 技术的优势

和其他的宽带接入技术相比,xDSL 技术的优势主要表现在以下几方面。

(1) 能够提供足够的带宽以满足人们对于多媒体网络应用的需求;

(2) 性能和可靠性有明显的优势;

(3) 利用现有的电话铜双绞线,能够平滑地与人们现有的网络进行连接,是比较经济的接入方案之一。

3. xDSL 技术的分类

xDSL 技术按上行(用户到交换局)和下行(交换局到用户)的速率是否相同可分为速率对称型和速率非对称型两种。根据信号传输的速率、距离,以及上行速率与下行速率的不同,主要的数字用户线 xDSL 技术可以分为以下几种。

(1) 非对称数字用户线(Asymmetric Digital Subscriber Line,ADSL);

(2) 高比特率数字用户线(High-bit-rate DSL,HDSL);

(3) 速率自适应数字用户线(Rate Adaptive DSL,RADSL);

(4) 甚高比特率数字用户线(Very High-bit-rate DSL,VDSL)。

表2-5给出了主要的 xDSL 技术的上行与下行速率的参数。

表2-5 主要的 xDSL 技术的参数

xDSL	下/上行速率(距离5.5km)	下/上行速率(距离3.6km)	线对数/对
ADSL	1.5Mbps/64kbps	6Mbps /640kbps	1
HDSL	1.544Mbps(对称)	1.544Mbps(对称)	2
VDSL	51Mbps/2.3Mbps	51Mbps/2.3Mbps	2
RADSL	1.5Mbps/64kbps	6Mbps/640kbps	1

数字用户线缩写 xDSL 中 x 的意思是表示它的不同类型,例如,可以理解 x 是 A、H、V 或 RA 等,它们对应于不同的数字用户线技术。

问题2-34:什么是 ADSL 与 ADSL-Lite 技术?

1. 非对称数字用户线 ADSL

非对称数字用户线 ADSL 技术最初是由 Intel、Compaq Computer、Microsoft 成立的特

别兴趣组(Special Interest Group,SIG)提出,如今这一组织已经包括大多数主要的 ADSL 设备制造商和网络运营商。

ADSL 主要的技术特点表现在以下几个方面。

(1) 它可以在现有的用户电话线上通过传统的电话交换网 PSTN,以重叠和不干扰传统模拟电话业务,同时提供高速数字业务。因此,ADSL 允许用户保留他们已经申请的模拟电话业务,可以同时支持单对用户电话线上的新型数据业务。新型的数据业务可以是 Internet 在线访问、远程办公、视频点播 VOD 等。

(2) 该技术几乎和本地环路的实际参数没有什么关系,因此与所使用的用户电话线的特性无关,因此用户不需要专门为获得 ADSL 服务而重新铺设电缆。

(3) ADSL 技术提供的非对称带宽特性,上行速率在 64～640kbps,下行速率在 500kbps～7Mbps。用户可以根据需要选择上行和下行速率。

这些特点对于网络运营商来说是很重要的,因为它意味着他们在推广 ADSL 技术时,用户端的投资相对比较小并且推广容易。

2. ADSL-Lite

传统的 ADSL 在用户住宅内安装无源分离器。无源分离器用来分离电话业务和数据业务。从无源分离器接出两条线,一条是普通的室内电话线路,用于普通的电话业务;另一条线接到住宅内的 ADSL Modem 上。ADSL Modem 一般放置在微型计算机旁边。安装无源分离器需要网络运营商派出技术工人来完成,这无形中增加了推广和维护的成本。

为了使 ADSL 的设备与线路安装更为简单,人们提出了一种新版本 ADSL 技术,称为 ADSL-Lite(轻型 ADSL)。ADSL-Lite 不需要安装无源分离器,电话和数据业务可以共享屋内的同一对双绞线。安装 ADSL Modem 就像安装普通的 Modem 一样简单。ADSL-Lite 又可以称作“简易经济型非对称数字用户线”。目前,大多数网络运营商向家庭用户推广的是符合 ADSL-Lite 标准的 Modem。ADSL-Lite 下行速率为 64kbps～1.5Mbps,上行速率为 32～512kbps。

问题 2-35:术语辨析:ISP、NAP、NSP、ICP 与 IDC。

1. 互联网服务提供商

(1) 我国信息产业部对接入服务有明确界定,将它作为“电信业务的第二类增值电信业务”。按照我国管理部门的界定,ISP(Internet Service Provider)是指利用接入服务器和相应的软硬件资源建立业务结点,并利用公用电信基础设施与 Internet 骨干网相连接,为各类用户提供 Internet 接入服务的企业。用户可以利用公用电话网、移动通信网、有线电视网或其他接入方式,通过 ISP 接入 Internet。

(2) Internet 接入服务业务主要有两种:一是为 ICP 经营者提供 Internet 接入服务;二是为普通上网用户提供 Internet 接入服务。

(3) ISP 是经国家主管部门批准的正式运营企业,受到国家法律的保护。这句话的潜台词是:攻击 ISP 是犯法的。

2. 国家服务提供商与网络接入点

第一层的国际或国家的 ISP 叫作国家服务提供商(National Service Provider,NSP)。第一层的国际或国家服务提供商 NSP 负责建设与维护主干网的公司,它们之间通过网络接入点(Network Access Point,NAP)互联。

3. 互联网内容提供商

(1) ICP(Internet Content Provider)是向网络用户提供信息服务和增值业务的企业。ICP同样是要经国家主管部门批准的正式运营企业,受法律保护。

(2) 根据主营业务划分,ICP类型主要有搜索引擎ICP、即时通信ICP、移动互联网业务ICP,以及提供新闻信息、文化信息等信息服务的门户ICP。

4. 互联网数据中心

(1) IDC(Internet Data Center)是在Internet的基础上建立标准化的、能够提供电信专业级机房环境,为企业、政府提供服务器托管、租用,以及数据存储服务的企业。

(2) IDC为ICP、企业、媒体和各类网站,提供高质量、安全可靠的专业化服务器托管、空间租用、网络带宽租用服务。

(3) IDC需要向租用方签订服务品质协议,规定双方的责任和权利,如果IDC服务违约将做出相应的经济赔偿。

(4) 提供云计算服务的企业就是一个典型的IDC。

第三部分　习题参考答案

1. 最大数据传输速率为6kbps。
2. 最大数据传输速率约为40kbps。
3. 第三个波长为33.3cm对应的频率为900MHz,不在ISM频段内。
4. (1) 二进制编码:01001011。

   (2) 差分曼彻斯特编码波形如图2-24所示。

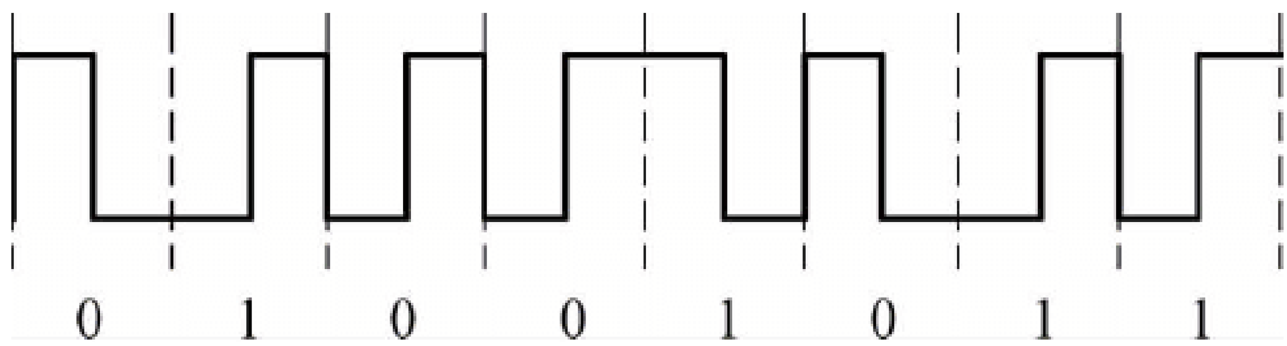


图2-24　差分曼彻斯特编码波形

5. 根据QAM调制中波特率与相数,计算出的对应比特率值如表2-6所示。

表2-6　调制速率与数据传输速率对应表

调制速率(baud)	多相调制的相数	数据传输速率(bps)
3600	QPSK-8	10 800
3600	QPSK-16	14 400
3600	QPSK-64	21 600
3600	QPSK-256	28 800

6. 可以容纳40条信道。
7. 传输信号的信道带宽至少为200MHz。
8. STM-4速率为622.08Mbps。
9. (1) A站没有发送;

   (2) B站发送0;

   (3) C站发送1;

   (4) D站发送1。

第 3 章 数据链路层

第一部分　学习目的、要求与知识点结构

1. 学习目的

在讨论了基于点-点通信线路的物理层协议与标准之后，本章将进一步讨论基于点-点通信线路的数据链路层协议与标准。本章将从差错产生的原因与差错控制方法入手，研究数据链路层的基本概念和服务功能，典型的数据链路层协议，以及数据链路层滑动窗口协议与帧传输效率分析方法。

2. 学习要求

(1) 理解：数据传输过程中差错产生的原因与性质。

(2) 掌握：误码率的定义与差错控制方法。

(3) 掌握：数据链路层的基本概念。

(4) 掌握：典型面向比特型数据链路层协议 HDLC 与 PPP。

(5) 掌握：数据链路层滑动窗口与帧传输效率分析方法。

3. 本章知识点的组织与结构

本章知识点的组织与结构如图 3-1 所示。

在图 3-1 中需要注意的几个问题如下。

(1) 第 3 章存在着从广度优先转入深度优先的变化，因此在本章内容的讨论中本着提出问题、分析问题、解决问题的思路，安排教学内容。从物理传输线路存在的传输差错问题出发，提出解决传输差错问题的差错控制机制，进而讨论实现差错控制机制的数据链路层协议。从数据链路层协议的分类入手，从入门级的面向字符型数据链路层协议，向面向字节型数据链路层协议逐步过渡，讨论了具有代表性的 HDLC 协议，然后重点分析 PPP，以及滑动窗口协议与效率分析方法。

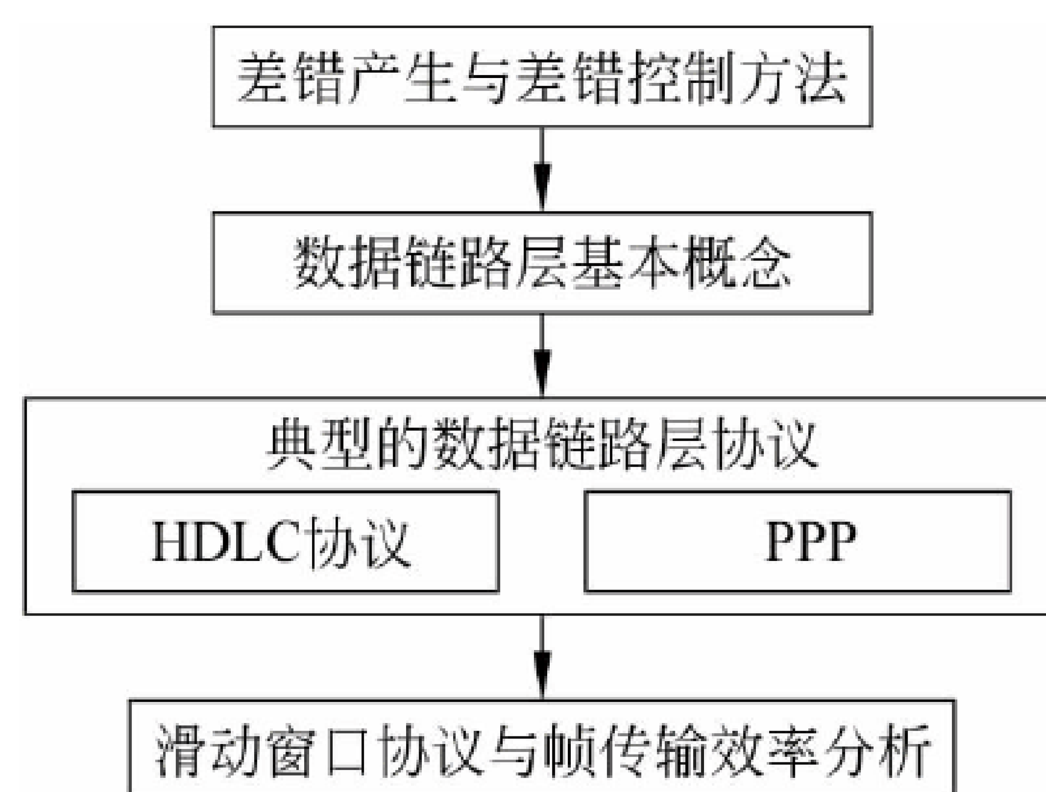


图 3-1　第 3 章的知识点结构

(2) 很多教材在讨论数据链路层时将面对点-点线路与面向广播线路的数据链路层协议放在一起讨论，对于网络初学者来说，很难接受两种完全不同类型的协议，经常在学完课

程复习的时候提出问题,恰恰证明是分不清楚两者通信环境的不同而引起的。所以作者从多年的教学经验出发,将它们分成两章讨论,分而治之,各有侧重。面向广播线路的数据链路层协议涉及的主要是局域网的问题。局域网 MAC 层的协议内容十分丰富,是读者学习网络技术的重要基础。我们在实际工作和生活中不可能不遇到 Ethernet 的问题,高速 Ethernet 如 GE、10GbE、40GbE、100GbE 技术目前正在发展,并且会从局域网扩展到城域网、广域网。另外,802.11 标准的无线局域网 Wi-Fi 目前已经广泛应用于办公室、校园、家庭。因此专门设立一章,重点展开讨论局域网技术是非常必要的。实际教学效果也证明了这一点。

(3) 一般理解 HDLC 协议与 PPP 的帧结构是相同的,应该都属于面向比特型数据链路层协议,而且 PPP 是 HDLC 协议的一个子集,因此 PPP 与面向字符型数据链路层协议无关。

但是,我们仔细研究 PPP 时会发现,PPP 与理想的 HDLC 协议有两点不同:HDLC 是典型的面向比特型的数据链路层协议,而 PPP 既可以用于同步通信,也可以用于异步通信之中,因此它需要具备面向比特型协议的特点,同时有能够处理面向字符型协议的能力。为了保证帧传输的透明性,当 PPP 用于同步通信时,使用 0 比特插入/删除方法;而用于异步通信时采用加入转义字符的方法。

第二部分 教学内容问答

问题 3-1:产生传输差错的主要原因是什么?

(1) 由于通信信道总是有一定的噪声存在,通信信道的噪声分为两类:热噪声和冲击噪声。当数据从信源出发,经过通信信道时,在到达信宿时,接收信号是数据信号与噪声的叠加。在接收端,接收电路在取样时判断信号电平。如果噪声对信号叠加的结果,在电平判决时引起判断失误,那么就会产生传输数据的错误。

(2) 传输误差是由热噪声和冲击噪声共同作用的结果。热噪声是由传输介质导体的电子热运动产生的。热噪声的特点是:时刻存在、幅度较小、强度与频率无关、频谱很宽、随机噪声。由热噪声引起的差错是一类随机差错。冲击噪声是由外界电磁干扰引起的。与热噪声相比,冲击噪声幅度较大,是引起传输差错的主要原因。

(3) 冲击噪声持续时间与数据传输中每比特的发送时间 T 相比,可能比较长,因而冲击噪声引起的相邻多个数据位出错呈突发性。冲击噪声引起的传输差错为突发差错。

问题 3-2:如何理解误码率的定义?

理解误码率的定义需要注意以下几个问题。

(1) 误码率是指二进制比特在数据传输系统中被传错的概率,它在数值上近似等于:

$$\mathrm{Pe} = \mathrm{Ne}/N$$

其中,N 为传输的二进制比特总数,Ne 为被传错的比特数。

(2) 误码率应该是衡量数据传输系统正常工作状态下传输可靠性的参数。

(3) 对于一个实际的数据传输系统,不能笼统地说误码率越低越好,要根据实际传输要求提出误码率要求;在数据传输速率确定之后,误码率越低,传输系统设备越复杂,造价越高。

(4) 对于实际数据传输系统,如果传输的不是二进制位,要折合成二进制位来计算。

(5) 差错的出现具有随机性,在实际测量数据传输系统时,只有被测量的传输二进制位数越大,才会越接近于真正的误码率值。

问题 3-3:检错码与纠错码的区别是什么?

回答这个问题需要注意以下几点。

(1) 在计算机通信中,能够自动检测出比特流在通信信道传输过程中产生的错误,并进行纠正的方法叫作差错控制方法。

(2) 差错控制首先要发现传输出错,再去纠正错误。这里有两种基本的模式:一是通过一种编码方法去发现错误,如何采取另一种机制去纠错;二是将查错与纠错同时进行。

(3) 对应两种模式就有两种编码方法:一是纠错码;二是检错码。

(4) 纠错码是在每个传输的帧头带上足够多的冗余信息,以便在接收端能够发现,并能够自动纠正传输出现的差错。

(5) 检错码是在每个传输的分组中带上一定的冗余信息,根据这些冗余信息,接收端可以发现出现了差错,但不能确定是哪一位或哪些位出错,并且自己不能纠正传输差错。

(6) 纠错码方法虽然有优越之处,但实现困难,在一般的通信场合不易采用。检错码方法虽然需要通过重传机制达到纠错的目的,但原理简单,实现容易,编码与解码速度快,目前正得到广泛的使用。

问题 3-4:如何理解循环冗余编码 CRC 的基本工作原理?

循环冗余编码(CRC)是目前应用最广的检错码编码方法之一,它具有检错能力强与实现容易的特点。

(1) CRC 检错方法的工作原理如下。

① 在数据链路层的协议中为发送端与接收端使用相同的生成多项式 $G(x)$。

② 在发送端用生成多项式 $G(x)$ 去除数据比特序列多项式,求得一个余数多项式。将余数多项式加到数据多项式之后发送到接收端。

③ 在接收端用同样的生成多项式 $G(x)$ 去除接收数据多项式 $f(x)$,得到计算余数多项式。如果计算余数多项式与接收余数多项式相同,则表示传输无差错;如果计算余数多项式与接收余数多项式不相同,则表示传输有差错,由发送方重发数据,直至正确为止。

实际的 CRC 校验码生成是采用二进制的模二算法(即减法不借位、加法不进位)计算出来的,这是一种异或操作。可以通过一些例子来形象地解释 CRC 的基本工作原理。

(2) 在用二进制的模二算法生成 CRC 校验码时,需要注意以下几个问题。

① 以 CRC-12 为例,$G(x)=x^{12}+x^{11}+x^3+x^2+x+1$,可以写为

$$G(x)=1\times x^{12}+1\times x^{11}+0\times x^{10}+0\times x^9+0\times x^8+0\times x^7+0\times x^6+0\times x^5+0\times x^4+1\times x^3+1\times x^2+1\times x+1\times x^0$$

尽管 CRC-12 的最高位是 x^{12},$k=12$。而实际上用二进制表示时,它的位数 $N=13$,也就是说用二进制表示 $G(x)$ 应该是:1100000001111。$k=N-1=13-1=12$。

② 如果在例子中给出生成多项式比特序列为 11001,那么写成生成多项式应该为

$$G(x)=1\times x^4+1\times x^3+0\times x^2+0\times x^1+1\times x^0$$

生成多项式的 $N=5$,$k=5-1=4$。

③ 从计算过程中可以看出,取 $k=4$ 实际上是将二进制比特序列左移4位,用这个4位放4位的余数。余数只能是等于或小于4位。因为按照二进制模二算法,如果还是5位,那么还可以继续去除。

问题3-5:为什么从教材给出的例子中看不出CRC能够检查出全部奇数位错?

(1) CRC检错效果。

① 能检查出全部单个错。

② 能检查出全部离散的两位错。

③ 能检查出全部奇数位错。

④ 能检查出全部长度小于或等于 K 位的突发错。

⑤ 能以 $[1-(1/2)^{K-1}]$ 的概率检查出长度为 $(K+1)$ 位的突发错。

(2) 有的读者针对书上的例子,发现如果出现奇数位错时,可能CRC校验不能够发现。这一点需要说明的是:CRC的检错能力是根据标准的生成多项式 $G(x)$ 计算出来的,由于在例子中使用的生成多项式没有经过严格的筛选,只是简单地表示CRC计算过程,因此达不到理论分析的检错能力。

问题3-6:如何理解差错控制机制的基本概念?

理解这个问题需要注意以下几点。

(1) 数据链路层的差错控制机制主要是通过反馈重发机制来实现的。

(2) 反馈重发机制的基本工作原理是:接收端通过检错码检查传送一帧数据是否出错,一旦发现传输错误,则通常采用反馈重发(ARQ)方法来纠正。反馈重发纠错实现方法有两种:停止等待方式和连续工作方式。

(3) 在停止等待方式中,发送方在发送完一帧后,要等待接收方的应答帧的到来。如果应答帧表示上一帧已正确接收,发送方就可以发送下一数据帧,否则重发出错数据帧。

(4) 连续ARQ协议的方法有两种:拉回方式和选择重发方式。

(5) 在拉回方式中,发送方可以连续向接收方发送数据帧,接收方对接收的数据帧进行校验,然后向发送方发回应答帧。如果发送方在连续发送了编号为 n 的帧后,从应答帧得知第 n 号数据帧传输错误,那么发送方将停止当前数据帧的发送,重发 n 号及之前的帧。

(6) 选择重发方式与拉回方式的不同之处在于:如果在发送完编号为 n 的数据帧时,接收到编号为 k 的数据帧传输出错的应答帧,那么发送方只重发出错的 k 号数据帧。

问题3-7:物理线路与数据链路是什么关系?

理解物理线路与数据链路的区别与联系,需要注意以下几个问题。

(1) 物理线路是由传输介质与通信设备构成的。以频带传输为例,连接收发双方的传输介质是电话线。由于电话线是用来传输模拟语音信号的,在电话线上传输计算机产生的数字信号就必须使用调制解调器Modem,实现数字信号与模拟信号之间的转换。收发双方的物理层通过电话线与Modem完成比特流的传输。因此,电话线与Modem就构成了连接收发双方物理层、实现比特流传输的物理线路。

(2) 没有采取差错控制机制的物理线路传输比特流会出错。在计算机网络中,设计数据链路层的目的是为了发现和纠正物理线路传输过程中的差错问题,使有差错的物理线路变成无差错的数据链路。数据链路是由实现数据链路层协议的硬件、软件与物理线路构成。

(3) 物理线路的比特流传输功能是由实际存在传输介质与通信设备实现的,而数据链路是逻辑上存在的,它的功能是通过数据链路的协议数据单元的帧头,按照数据链路层协议规定的协议动作来实现的。

(4) 在实际工作过程中,首先要建立物理线路连接,在物理线路上能够传输比特流之后,才有可能传输数据链路建立的协议帧,建立数据链路,传输数据帧。在通信结束之后,先释放数据链路连接,之后才能释放物理线路连接。

问题 3-8:数据链路的主要功能是什么?

数据链路的控制功能是为了保证数据通过物理线路传输的正确性。数据链路控制的主要功能有以下几点。

1. 链路管理

(1) 当两个结点要开始进行通信时,发送方必须确知接收方是处在准备接收数据的状态。为此,双方必须先交换一些必要的信息,建立数据链路连接。

(2) 在传输数据时要维持数据链路。当通信完毕时要释放数据链路。

(3) 数据链路的建立、维持和释放就叫作链路管理。

2. 帧同步

(1) 在数据链路层,数据以帧为单位传送。物理层的比特流按照数据链路层协议的规定被封装在数据帧中传送。

(2) 帧同步是指收方应当能从收到的比特流中准确地区分出一帧的开始和结束在什么地方。

3. 流量控制

发送方发送的数据必须使接收方来得及接收。当接收方来不及接收时,就必须控制发送方发送数据的速率。

4. 差错控制

计算机通信要求必须保证极低的误码率,为此必须采用差错控制技术。

差错控制技术要使接收端能够发现传输错误,并能纠正传输错误。数据链路层实体将对帧的传输过程进行检查,发现差错并能用反馈重发纠错等方法纠正传输错误。

5. 透明传输

当所传的数据中出现了控制字符时,就必须采取适当的措施,使接收方不至于将数据误认为是控制信息,这样才能保证数据链路层的传输是透明的。

6. 寻址

在多点连接的情况下,要保证每一帧能传送到正确的目的结点。接收方也应当知道发送方是哪一个结点。

问题 3-9:数据链路层协议有哪几种类型?

(1) 数据链路层协议基本可以分为两类:面向字符型与面向比特型。

(2) 数据链路层协议经历了一个不断改进的过程。最早出现的数据链路层协议是面向字符型的协议。它的特点是利用已定义好的一种标准字编码(如 ACSII 码或 EBCDIC 码)的一个子集来执行通信控制功能。面向字符型协议可以利用 ASCII 码中的 10 个控制字符(例如报头开始 SOH、肯定应答 ACK、同步字符 SYN 等)实现通信控制功能,典型的面向字

符型数据链路层协议是二进制同步通信(Binary Synchronous Communication,BSC)协议。

(3) 面向字符型协议有两个明显缺点:一是使用不同字符集的计算机的控制字符不一样,因此会给通信造成困难。在面向字符型的BSC协议中,使用ASCII码中的10个控制字符完成通信控制功能,并规定了数据与控制报文的格式,以及协议操作过程。使用不同字符集的两台机器很难利用面向字符型协议进行通信。

二是控制字符的编码(例如同步字符SYN编码为0010110)不能在用户数据字段中出现。如果数据字段中出现与控制字符相同的编码时,就会引起通信控制错误。

(4) 为了克服这两个主要缺点,在此基础上提出面向比特型协议。典型的面向比特型协议主要有:高级数据链路控制HDLC协议与PPP。

问题3-10:为什么说系统学习HDLC协议对理解数据链路层协议有重要的作用?

根据作者的科研和教学经验,理解HDLC对于掌握计算机网络协议的设计方法是很重要的。第一步迈出去了,之后的数据链路层协议,甚至是网络层协议都有很多相通之处。很多网络应用软件的编程都会运用HDLC的设计思想。同时,很多数据链路层协议取的是HDLC的子集,例如,PPP取的是HDLC的子集,Ethernet帧结构是在HDLC帧结构基础上演变而来的。因此,HDLC应该是理解网络工作原理的入门级协议。HDLC协议的内容相对比较复杂,理清它的内在关系不太容易,但是它比起网络层的IP协议还是简单的。很多教科书在这一点上都采取了粗线条的处理方法,其实不是上策。

作者希望通过HDLC协议的学习,引导读者深入到内部细节上去了解网络协议的研究、制定与实现的方法。因此,作者建议在介绍HDLC协议之前,首先让学生了解如图3-2所示的数据链路配置和数据传输方式的基本关系。实际上这种结构是协议制定者,根据需要解决问题的实际环境来确定的,作者考虑了可能出现的各种情况。建立起整体的概念之后,再讨论滑动窗口与协议效率问题就会顺畅得多。

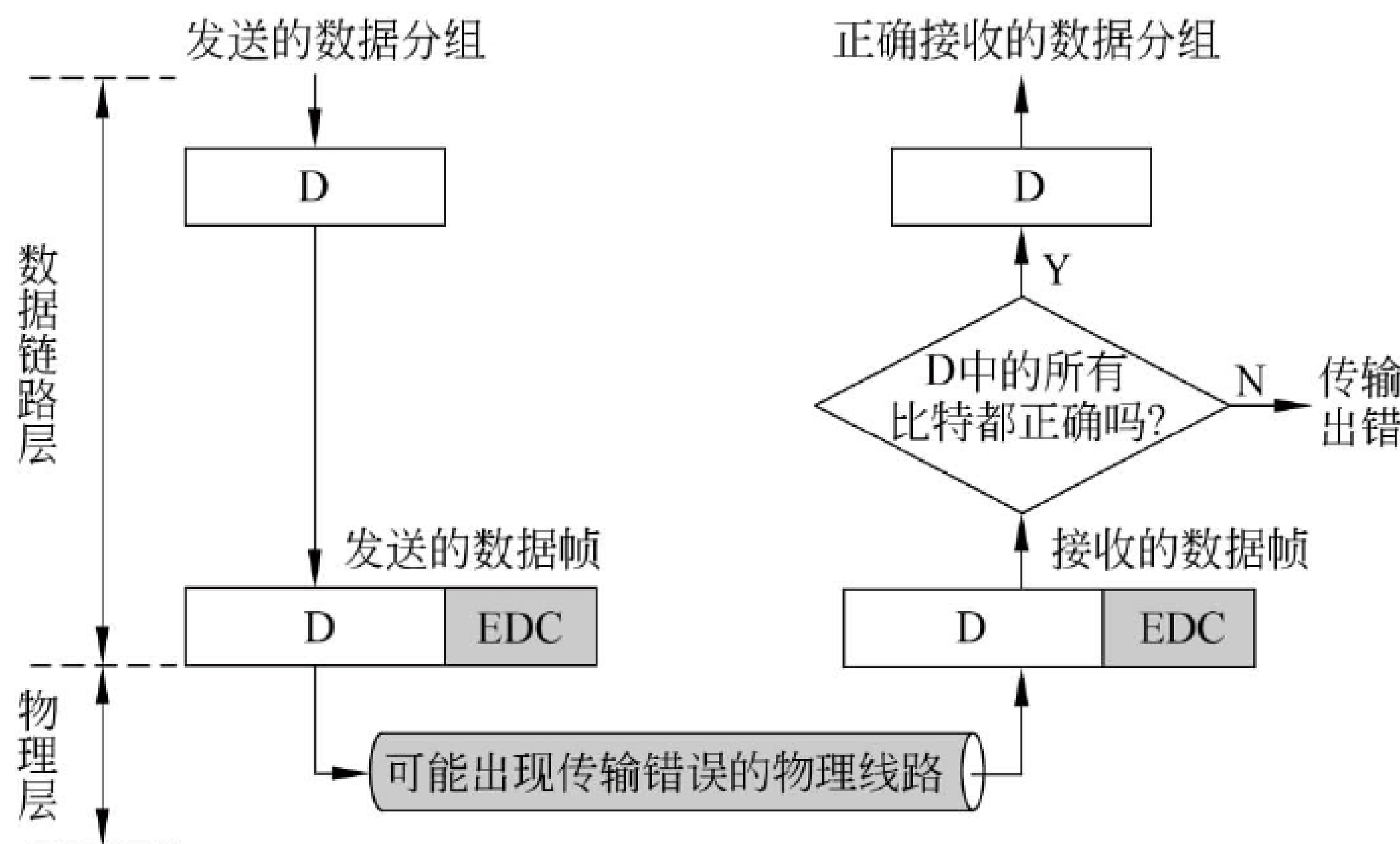


图3-2 数据链路层协议在保证传输可靠性方面的作用

随着计算机通信的发展,这种面向字符型链路控制规程逐渐暴露出缺点。针对面向字符型协议的缺点,人们认识到必须设计出一种新的数据链路层协议来代替旧的面向字符型协议。

1974年,IBM公司推出了SNA体系结构,在数据链路层采用了面向比特型SDLC协议。IBM建议美国国家标准协会ANSI和国际标准化组织ISO将SDLC协议变成国际标

准。ANSI将SDLC修改后的先进数据通信规程(Advanced Data Communication Control Procedure,ADCCP)作为美国国家标准。ISO将SDLC修改后的高级数据链路控制(High-level Data Link Control,HDLC)协议作为国际标准ISO 3309。

问题3-11:如何理解HDLC对数据链路的配置方式和数据传送方式?

初次接触HDLC协议的读者经常对数据链路配置和数据传输方式搞不清楚。图3-3给出了HDLC协议数据链路配置和数据传输方式的示意图。

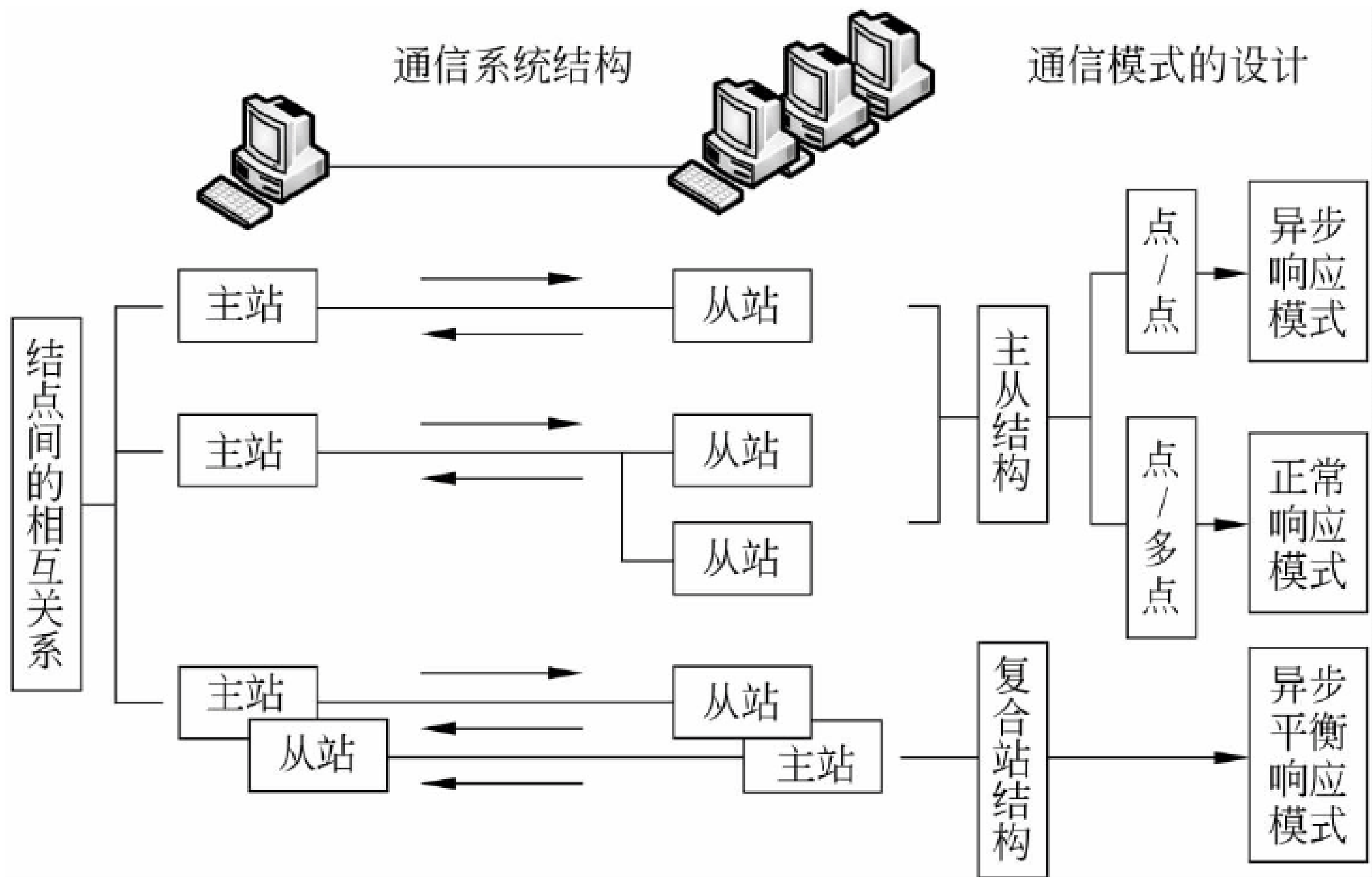


图3-3 HDLC协议数据链路配置和数据传输方式的示意图

HDLC数据链路有两种基本配置方式:非平衡配置与平衡配置。

1. 非平衡配置方式

(1) 主站与从站的结构。

非平衡配置的特点是:一组结点根据在通信过程中的地位分为主站与从站,由主站来控制数据链路的工作过程。主站发出命令;从站接收命令,发出响应,配合主站工作。

点-多点方式的链路中,主站可以与每个从站之间分别建立数据链路。

(2) 正常响应模式与异步响应模式。

非平衡配置可以有两种数据传送方式:正常响应模式(NRM)与异步响应模式(ARM)。

在正常响应模式中,主站可以随时向从站发送数据帧。从站只有在主站向它发送命令帧探询、从站响应后,才可以向主站发送数据帧。

在异步响应模式中,主站和从站可以随时相互发送数据帧。从站不需要等待主站发出探询就可以发送数据帧。但是,主站仍然负责数据链路的初始化、链路的建立、释放与差错恢复等功能。

2. 平衡配置方式

(1) 平衡配置的特点是链路两端的两个站都是复合站。

(2) 复合站同时具有主站与从站的功能,每个复合站都可以发出命令与响应。

(3) 平衡配置结构只有一种工作模式——异步平衡模式(ABM),每个复合站都可以平等地发起数据传输,而不需要得到对方复合站的许可。

实际上,这是协议的研究者在设计这样一个点-点与点-多点结构的数据通信模式时,设想的各种可能的连接方式与通信方式。

问题 3-12:如何理解 HDLC 帧结构特点?

理解 HDLC 帧结构特点时需要注意以下几点。

(1) 帧结构组成为:标志字段 F、地址字段 A、控制字段 C、信息字段 I、帧校验字段 FCS。

(2) HDLC 帧分为三大类:信息帧 I(Information)、监控帧 S(Supervisory)与无编号帧 U(Unnumbered)。

(3) 帧校验字段 FCS 采用 CRC 校验方式,校验范围为 A、C、I 字段。

(4) 信息帧的发送序号 N(S)表示当前发送的信息帧的序号,接收序号 N(R)表示已正确接收序号 N(R)-1 及以前的各帧,具有捎带确认的作用。

(5) 探询/终止(P/F)位等于 1,在帧交换的过程中成对出现。

(6) 监控帧(S 帧)用在选择重传机制中。

(7) 无编号帧(U 帧)用于表示:置 SARM、SNRM 或 SABM 模式,以及拆链 DISC 的命令与响应。

(8) HDLC 解决帧透明性的方法是:0 比特插入/删除。

问题 3-13:如何理解 HDLC 协议的交互过程?

在教学过程中,可以以主教材中图 3-4 给出了正常响应模式的工作过程为例,说明 HDLC 建立数据链接、传输数据帧、释放数据连接的协议执行过程。这个例子尽管简单,但是具有代表性。读者如果能够理解这样一个简单的协议交互过程,对于以后学习中比较复杂的过程接受起来相对就会容易一些了。

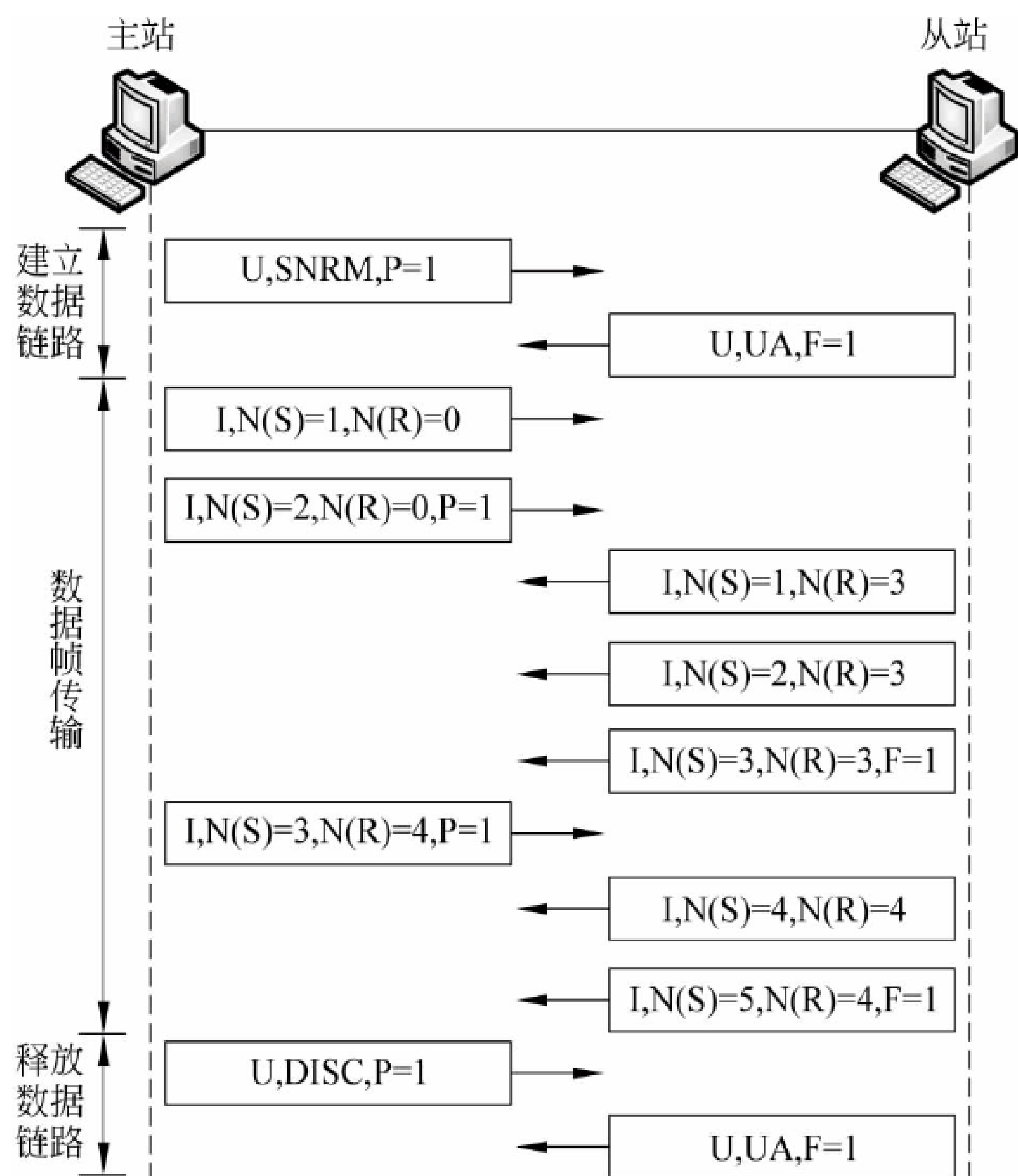


图 3-4 正常响应模式的工作过程示意图

正常响应模式的工作过程分为以下三个阶段。

1. 建立数据链路阶段

当两个结点要在正常响应模式之下建立数据链路连接，由主站使用置正常响应模式SNRM帧，从站用无编号确认UA帧向主站应答。

2. 数据帧传输阶段

在正常响应模式中，主站可以随时向从站传输数据帧；从站只有在主站发送命令帧中探询位P=1进行探询，从站用F=1响应后才能向主站发送数据帧。

在帧传输过程中，主站与从站通过N(S)表示下一个要发送的帧序号，用N(R)进行捎带确认。

3. 释放数据链路阶段

进入释放数据链路连接的阶段时，主站发送拆链命令DISC帧，用"U、DISC、P=1"表示。当从站同意释放数据链路时，它将用无编号确认UA帧向主站应答，用"U、UA、P=1"表示。当主站接收到UA帧时，释放数据链路。

从以上过程的讨论中可以知道，在HDLC协议的作用下，收发双方可以有条不紊、正确地完成数据帧传输，使物理线路上可能出现的数据传输错误得到及时发现与纠正，从而提高数据传输的可靠性。

问题3-14：数据链路层滑动窗口协议有哪几种类型？

数据链路层的差错控制与流量控制采用了滑动窗口协议(Sliding Windows Protocol)。数据链路层滑动窗口控制协议分为单帧停止等待协议与多帧连续发送协议。

单帧停止等待反馈重发纠错协议简称为单帧停止等待协议或一位滑动窗口协议。

多帧连续发送协议又可以进一步分为后退N帧(GBN)协议与选择重传(SR)协议。后退N帧协议也称为拉回重传协议。

滑动窗口协议的控制策略对数据链路层帧传输效率有很大的影响。

问题3-15：如何分析单帧停止等待协议的效率？

图3-5给出了单帧停止等待协议的帧传输过程示意图。

1. 分析的前提

(1) 理想状态下帧传输总延时的计算公式。

(2) 理想状态是指一个数据帧从发送方正确传输到接收方，接收方返回的确认帧也正确传输到发送方，在一次数据帧与确认帧的传输中没有出现错误。

2. 理想状态下帧传输总延时分析结果

$$t_T = t_p + t_f + t_{pr} + t_a + t_p + t_{pr} = 2t_p + 2t_{pr} + t_f + t_a \qquad (3\text{-}1)$$

对式(3-1)可以做两点简化。

(1) 帧的处理延时t_{pr}小于帧发送延时t_f与传播延时t_p，t_{pr}可以忽略。

(2) 确认帧通常很短，确认帧ACK的发送延时t_a可以忽略。

简化后的帧传输总延时为

$$t_T \approx t_f + 2t_p \qquad (3\text{-}2)$$

在理想状态下，停止等待协议的帧传输效率为

$$U = t_f/(t_f + 2t_p) \qquad (3\text{-}3)$$

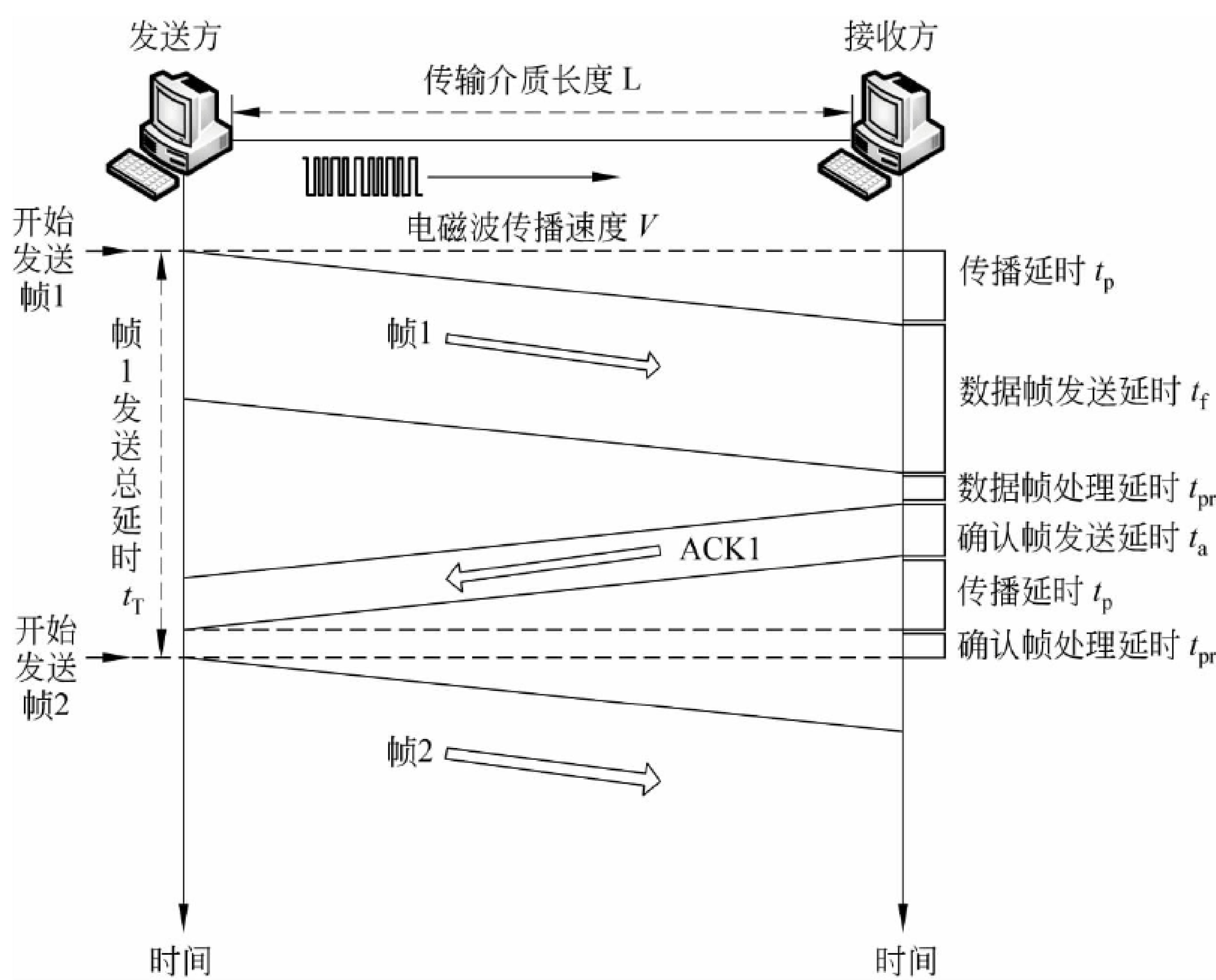


图 3-5　单帧停止等待方法的帧传输过程示意图

假设：α=传播延时/发送延时= t_p/t_f

$$U = 1/(1+2\alpha) \tag{3-4}$$

3. 讨论

从式(3-4)可以看出，数据链路层帧传输效率U受α的影响。这里可以做出这样两个假设：一是假设两个结点的距离一定，则传播延时t_p值也一定；二是假设结点发送速率一定。

我们通过以下参数的计算，来讨论影响协议效率的因素。

(1) 如果电磁波在有线传输介质(如电缆)中的传播速度约为空间电磁波传播速度的2/3，空间电磁波传播速度为3×10^8(m/s)，则在电缆中传播速度约为2×10^8(m/s)。如果连接收发双方的传输介质长度为1000m，则传输延时t_p约等于5×10^{-6}(s)。

(2) 如果数据帧长度为100b，结点的发送速率为10Mbps，则发送延时$t_f=1\times10^{-5}$(s)。

$$\alpha_1 = t_p/t_f = 5\times10^{-6}/(1\times10^{-5}) = 0.5$$

$$U_1 = 1/(1+2\times0.5) = 0.50$$

(3) 如果数据帧的长度为1000b，其他参数不变，那么发送延时$t_f=1\times10^{-4}$(s)。

$$\alpha_2 = t_p/t_f = 5\times10^{-6}/(1\times10^{-4}) = 0.05$$

$$U_2 = 1/(1+2\times0.05) \approx 0.91$$

(4) 比较U_1、U_2可以看出，当传播延时一定时，发送的数据帧长度越长，发送延时就越大，α就越小，传输效率U也就越高。发送的数据帧长度越短，发送延时就越小，α就越大，传输效率U就越低。

(5) 推论：在保持t_f+2t_p时间内不出现差错的条件下，连续发送多个帧，可以提高帧传输效率。

问题 3-16：如何理解滑动窗口控制机制的工作原理?

理解这个问题需要注意以下几点。

(1) 多帧连续发送协议包括后退 N 帧(GBR)的拉回重发方式与选择重发(SR)方式。选择重发 SR 的效率高于后退 N 帧 GBR。

(2) 在讨论数据链路层差错控制的 GBN 方式与 SR 方式中,发送方不必等待接收方的确认 ACK 信息到来,就可以连续发送多个数据帧。但是从流量控制的角度,发送方可以连续发送帧的数量要受到接收方的限制。限制的因素主要是:接收方的接收缓冲区可以用于接收新的接收帧的字节数,接收方处理数据帧的速度,以及接收方需要等待重传的帧有多少。HDLC 协议帧头发送序号 N(S)与接收序号 N(R)在差错控制中能够起到"捎带确认"的作用,同时也可以用于数据链路层流量控制的滑动窗口协议之中。

(3) 在滑动窗口协议中定义了发送窗口与接收窗口。滑动窗口通过协调发送窗口 Wr 值与接收窗口值 Ws 的方法来实现流量控制功能。

(4) 数据链路层协议采取滑动窗口控制时,在确定发送窗口与接收窗口大小、接收方在什么时间发送对哪个帧的确认时,都需要考虑到接收方的网络层能读取多少个接收帧、接收缓冲区可以拿出多少空间准备接收新的数据帧,以及接收方需要等待多少个重传帧等因素。

(5) 滑动窗口控制正是在考虑以上因素的基础上,调节和控制数据链路层发送方与接收方之间的数据流量。

问题 3-17:PPP 经历了怎样的发展过程?

回答这个问题需要注意以下几点。

(1) 人们将用于串行线路的 Internet 数据链路层协议(SLIP)与点-点协议(PPP)叫作 Internet 数据链路层协议,目前最流行的是 PPP。

(2) SLIP 与 PPP 主要用于串行通信的拨号线路上,是目前家庭计算机或公司用户通过 ISP 方式连接到 Internet 的主要协议。

(3) SLIP 的历史可以追溯到 20 世纪 80 年代初,它最早是在 BSD UNIX 4.2 版操作系统上实现的。SLIP 支持 TCP/IP,它只是对数据报进行了简单的封装,然后用 RS-232 接口串行线路进行传输。SLIP 通常也用来将远程终端连接到 UNIX 主机,也可通过租用或拨号串行线路进行主机到路由器,以及路由器到路由器的通信。

(4) PPP 是 SLIP 新的版本。PPP 代替了 SLIP,并解决了 SLIP 中的一些效率问题。PPP 是在大多数家庭个人计算机和 Internet 服务提供商 ISP 之间使用的协议,它在高速广域网上也有一定的应用。一些社区宽带网也将 PPP 作为自己协议族的一部分。

问题 3-18:PPP 具有哪些主要的特点?

PPP 的主要特点表现以下几个方面。

(1) 不使用帧序号,不提供流量控制功能。

(2) 只支持点-点连接,不支持点-多点连接。

(3) 只支持全双工通信,不支持单工与半双工通信。

(4) 可以支持异步、串行通信,也可以支持同步、并行传输。

问题 3-19:PPP 具有哪些基本的功能?

PPP 可以提供以下基本功能。

(1) 用于串行链路的基于 HDLC 数据帧的封装机制。

(2) 链路控制协议(LCP)用以建立、配置、管理和测试数据链路连接。

(3) 网络控制协议(NCP)用以配置不同的网络层协议。

需要注意的是:网络控制协议(NCP)应该不属于数据链路层的问题。

问题 3-20:PPP 协议帧有几种类型?

RFC1660、RFC1661 定义了 PPP 和帧结构。根据 PPP 的基本功能,PPP 帧分为三种类型:链路控制 LCP 帧、网络控制 NCP 帧与 PPP 信息帧。

从严格的网络层次结构概念出发,链路控制 LCP 帧用于建立、配置、管理和测试数据链路连接,PPP 信息帧用于传输网络层数据都可以理解,但是 PPP 多出了网络控制 NCP 帧。

设计网络控制 NCP 帧说明了两个问题:一是 PPP 制定的时间相对早一些;二是 PPP 应用于电话交换网、ADSL 与 HFC 传输网,以及连接路由器-路由器的数据链路上,尤其涉及 ISP 的动态 IP 地址分配,因此出现了网络控制 NCP 帧。

问题 3-21:如何理解 PPP 链路控制帧的作用?

理解 PPP 链路控制帧的作用需要注意以下几个问题。

(1) 一台个人计算机通过 ISP 成为 Internet 中一台临时主机的工作过程如下三步。

第 1 步:个人计算机通过调制解调器呼叫 ISP 的路由器。

第 2 步:在路由器的调制解调器回答了电话呼叫后,物理连接建立。

第 3 步:个人计算机发送给路由器的链路控制帧,这些链路控制帧可用来指定 PPP 数据链路的选项。

(2) 网络控制 NCP 帧 PPP 的数据链路选项如下。

① 链路控制帧可以用来与对方进行协商,异步链路中将什么字符当作转义字符。

② 为提高线路的利用率,链路控制帧可以用来与对方进行协商,是否可以不传输标志字节或地址字节,并且将协议字段从两个字节缩短为一个字节。

③ 如果在线路建立期间,收发双方不使用链路控制协商的话,固定的数据字段长度为 1500B。

问题 3-22:PPP 在解决帧透明性时有什么特殊之处?

PPP 在解决帧透明性时的特殊之处表现在:可以用于异步通信,也可以用于同步通信。

(1) 当 PPP 用于异步通信时,信息字段中不能出现与标志字段“0x7E”相同的值,这就是帧传输“透明性”问题。为了解决这个问题,RFC1662 定义转义字符“0x7D”,并且使用字节填充。

(2) 当 PPP 用于同步通信时,为了保证帧传输的透明性,采用“0 比特插入删除”方法。

第三部分　习题参考答案

1. 发送的比特序列:1110001111010

曼彻斯特编码序号波形图如图 3-6 所示。

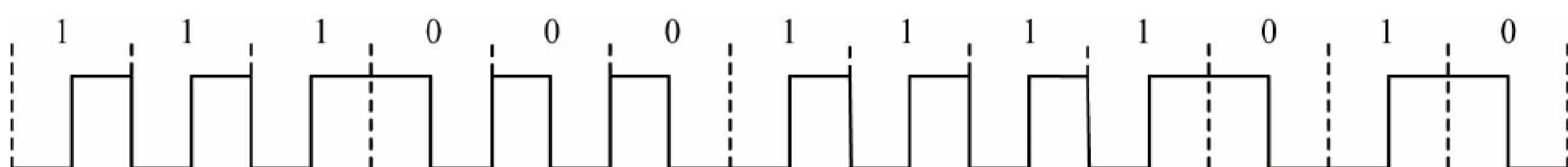


图 3-6　曼彻斯特编码序号波形图

2. 从 CRC 校验中可以发现传输出现差错。
3. 发送方需要重发编号为 1～6 的 6 个帧。
4. 发送方需要重发编号为 1、3 的两个帧。
5. (1) 停止-等待协议的信道最大利用率为 3.57%。
   (2) 连续传输协议的信道最大利用率为 12.90%。
6. 发送数据最少用 2×10^5(s)。

第 4 章 介质访问控制子层

第一部分　学习目的、要求与知识点结构

1. 学习目的

本章在介绍介质访问控制 MAC 子层关键技术与局域网体系结构、协议标准的基础上，对共享介质局域网、交换局域网、VLAN 与高速局域网的工作原理与组网方法，以及 WLAN 的工作原理进行系统的讨论。

通过本章的学习，读者能够了解局域网的主要技术特点，掌握主流局域网技术——Ethernet 的基本工作原理，高速 Ethernet——FE、GE、10GbE、40GbE、100GbE，以及交换局域网、VLAN 与 Wi-Fi 的基本工作原理，掌握局域网互联的基本概念和网桥的基本工作原理，初步具备局域网组网的基础知识与能力。

2. 学习要求

（1）了解：局域网的分类与特点。

（2）理解：IEEE 802 参考模型与介质访问控制子层的基本概念。

（3）掌握：Ethernet 局域网的基本工作原理。

（4）掌握：高速局域网、交换局域网与 VLAN 的基本工作原理。

（5）掌握：无线局域网 Wi-Fi 与 802.11 标准的基本概念。

（6）掌握：网络互联基本概念与网桥的基本工作原理。

3. 本章知识点的组织与结构

本章知识点的组织与结构如图 4-1 所示。

根据技术的发展与教材内容的安排，第 2 版将第 1 版的第 5 章中 Token Ring、Token Bus 这两种类型的局域网的原理讨论进行了简化和压缩，增加了 GE、10GbE 与 802.11 WLAN 的内容与分量；第 3 版在此基础上增加了 40GbE 与 100GbE 与 Wi-Fi 协议的相关内容。

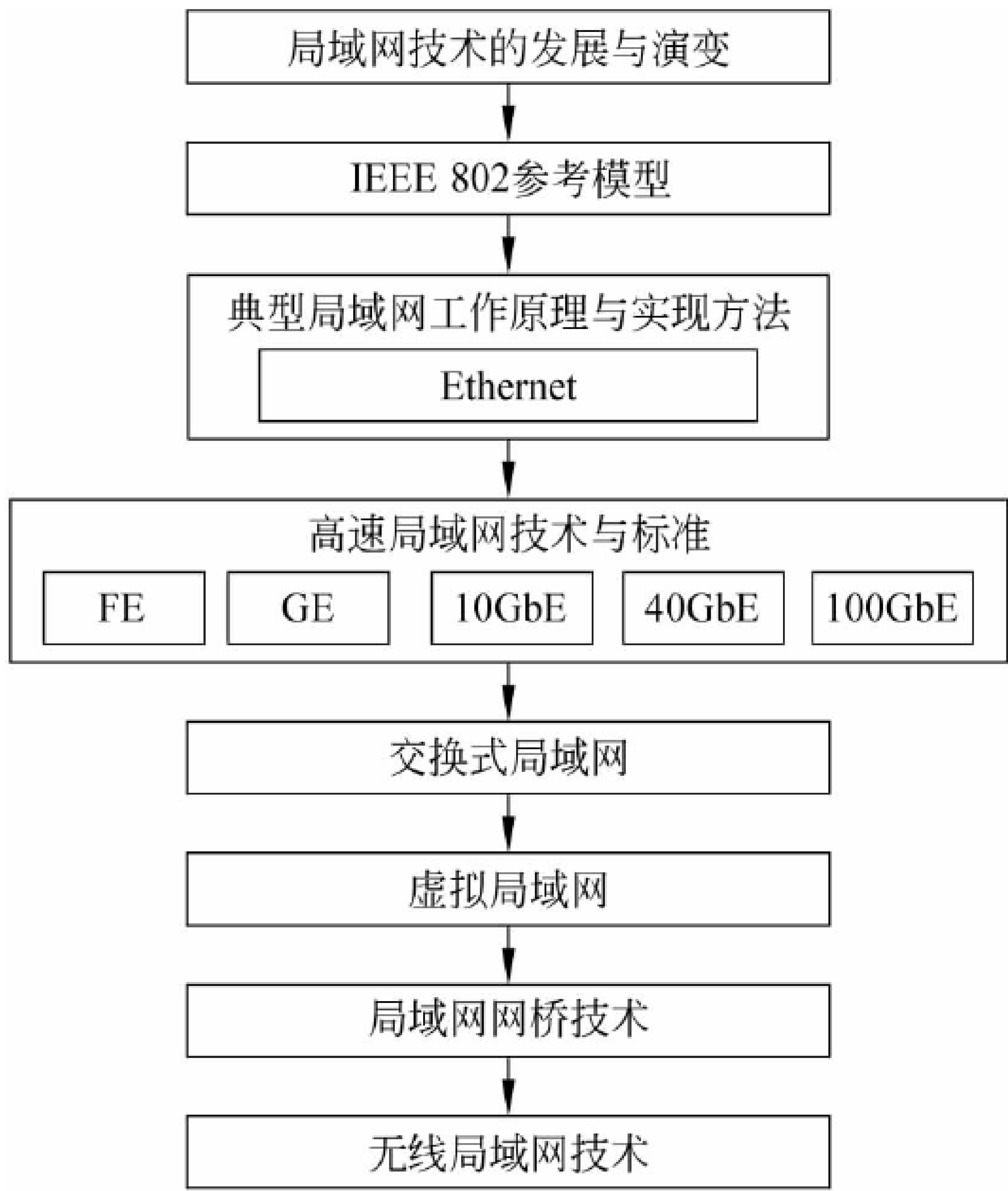


图 4-1　知识点的组织与结构

第二部分　教学内容问答

问题 4-1：如何认识局域网发展与演变的过程？

理解局域网发展的历史需要注意以下几个问题。

(1) 图 4-2 给出了局域网技术发展的过程示意图。图中按照时间的顺序列出了对局域网发展有影响的不同时期标志性的技术。

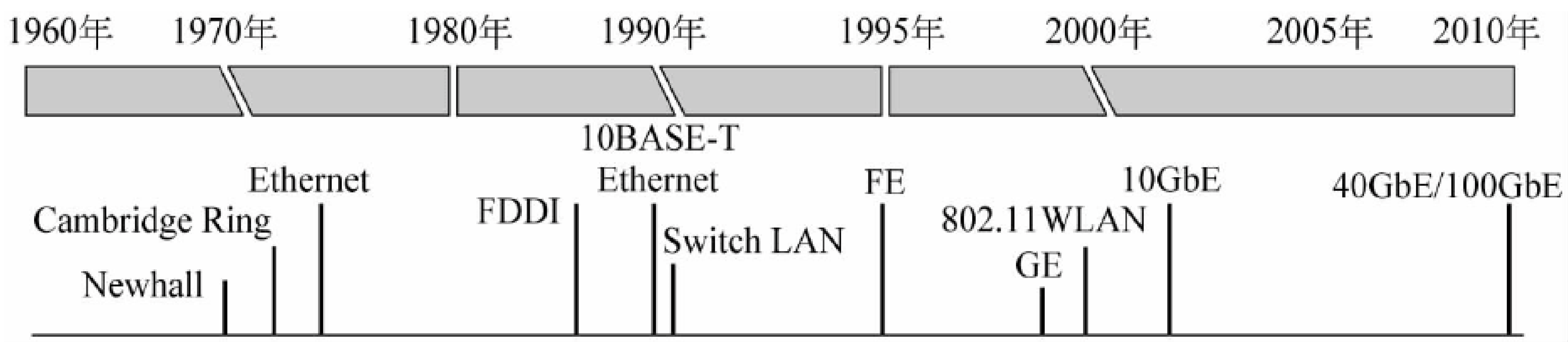


图 4-2　局域网技术发展的过程

(2) 从局域网发展的背景看，广域网技术的成熟与微型计算机的广泛应用，推动了局域网技术研究的发展。局域网是继广域网之后，网络研究与应用的又一个热点。20 世纪 80 年代，随着个人计算机技术的发展和广泛应用，用户共享数据、硬件与软件的愿望日益强烈。这种社会需求导致局域网技术出现了突破性的进展。

(3) 在局域网研究领域中，Ethernet 技术并不是最早，但它是最成功的技术。20 世纪 70 年代初期，欧美的一些大学和研究所开始研究局域网技术。早期的局域网主要是令牌环网。例如，1972 年美国加州大学研究 Newhall 环网，1974 年英国剑桥大学研究 Cambridge Ring 环网。这些研究成果对局域网技术的发展起到重要作用。20 世纪 80 年代，局域网领

域出现 Ethernet 与 Token Bus、Token Ring 三足鼎立的局面,并且各自都形成相应的国际标准。

(4) 1980 年 2 月 IEEE 召开了第一次“局域网标准会议”,研究速率为 1~20Mbps 的局域网标准。项目代号为“802”,因此之后制定的局域网协议标准都属于“IEEE 802 标准”系列。IEEE 803 委员会采用了施乐(Xerox)公司的 Ethernet DIX 标准为基础,1985 年发布了“IEEE 802.3 CSMA/CD 方法和物理层规范”。尽管施乐公司已经放弃了对以太网的商标所有权,但是 IEEE 仍然没有将“Ethernet”写入规范中。IEEE 将这项技术简称为“802.3 CSMA/CD”或“802.3”。

(5) 1990 年,IEEE 802.3 标准中的物理层标准 10BASE-T 推出,使非屏蔽双绞线可以作为 10Mbps 的 Ethernet 传输介质。在使用非屏蔽双绞线以后,Ethernet 组网的造价降低,可靠性提高,性能价格比大大提升,这就使 Ethernet 在与其他局域网的竞争中占据了明显优势。同年,Ethernet 交换机产品面世,标志着交换 Ethernet 的出现。

(6) 1993 年,Kalpana 研究了全双工 Ethernet,它改变了传统 Ethernet 的半双工工作模式,使得 Ethernet 带宽增加了一倍。在此基础上,利用光纤作为传输介质的物理层标准 10 BASE-F 和产品的推出,使得 Ethernet 技术最终从三足鼎立中脱颖而出。

开放的 Ethernet 技术与标准,使它得到软件开发商与硬件制造商的广泛支持。网络操作系统 NetWare、Windows NT Server、IBM LAN Server 及 UNIX 操作系统的应用,使得 Ethernet 技术进入了成熟阶段。到 20 世纪 90 年代,Ethernet 开始受到业界认可和广泛应用。21 世纪,Ethernet 技术已成为局域网领域的主流技术。

IEEE 802.3 标准由 IEEE 802.3 LAN/MAN 标准委员会(LMSC)负责维护。2012 年修改的 IEEE Std 802.3—2012 名称为“IEEE 以太网标准”,标准共有 3747 页。IEEE 提供免费下载的地址是 http://standards.ieee.org/about/get/802/802.3.thml。

问题 4-2:如何认识高速 Ethernet 的发展背景?

理解高速 Ethernet 发展的背景需要注意以下几个问题。

(1) 促进局域网发展的直接因素是个人计算机性能的提高与 Internet 的广泛应用。个人计算机的处理速度迅速上升,而价格却在很快下降,这进一步促进了个人计算机的广泛应用。大量用于办公自动化与信息处理的计算机必然要联网,这就造成局域网规模的不断增大和网络通信量的进一步增加。因此,局域网的带宽与性能已不能适应要求。各种新的应用不断提出,个人计算机已从初期简单的文字处理、信息管理等应用发展到分布式计算、多媒体应用,用户对局域网的带宽与性能也有了更高的要求。同时,新的基于 Web 的 Internet 应用也要求更高的带宽。这些因素促使人们研究高速局域网技术,希望通过提高局域网的带宽来改善局域网的性能,以适应各种新的应用环境的要求。

(2) 传统的局域网技术建立在“共享介质”的基础上,网中的所有结点共享一条共用的通信传输介质。介质访问控制方法用来保证每个结点都能“公平”地使用传输介质。在网络技术的讨论中,可以粗略做一个估算,如果 Ethernet 中有 N 个结点,那么每个结点平均能够得到的带宽为 $10/N$(Mbps)。显然,随着局域网规模的不断扩大,结点数 N 的不断增加,每个结点平均能分配到的带宽将越来越少。也就是说,当网络结点数 N 增大时,网络通信负荷加重,冲突和重发次数将大幅增长,网络线路的利用率急剧下降,网络传输延迟明显增加,网络服务质量将会显著下降。为了克服网络规模与网络性能之间的矛盾,人们提出了以下

三种可能的解决方案：提高速率、变共享为交换、互联（如图 4-3 所示）。

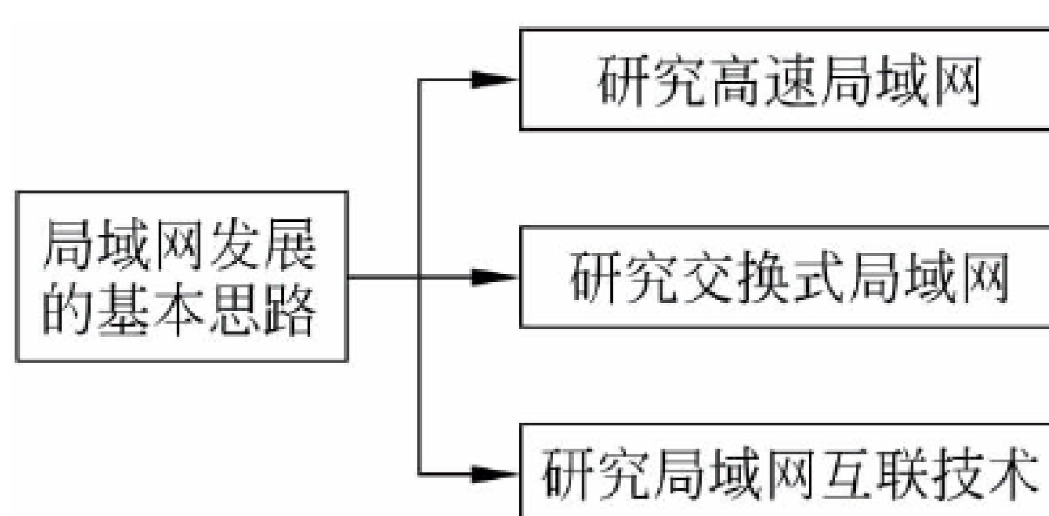


图 4-3　局域网发展的基本思路

（3）提高 Ethernet 的数据传输速率，从 10Mbps 提高到 100Mbps，甚至提高到 1Gbps 或 10Gbps、40Gbps 与 100Gbps，这就导致了高速局域网技术的研究。在这个方案中，无论局域网的传输速率提高到 100Mbps 还是 10Gbps，甚至是 40Gbps 与 100Gbps，它们保持 Ethernet 帧结构、最大与最小帧长度不变。

1995 年，传输速率为 100Mbps 的 Fast Ethernet 标准和产品推出。1998 年，传输速率为 1Gbps 的 Gigabit Ethernet 标准推出。1999 年，Gigabit Ethernet 的产品问世，并成为局域网主干网的首选方案。2002 年，数据传输速率为 10Gbps 的 Ethernet 标准正式通过。2010 年完成了数据传输速率为 40 Gbps 与 100 Gbps 的 Ethernet 标准的研究。这些都进一步增强了 Ethernet 在局域网应用中的竞争优势。在局域网工程领域中，人们经常将 Fast Ethernet、Gigabit Ethernet、10Gbps 的 Ethernet，以及速率为 40Gbps、100Gbps 的 Ethernet 简称为 FE、GE、10GbE、40GbE 与 100GbE，而将 10Mbps 的 Ethernet 简称为传统 Ethernet 或 Ethernet。Ethernet 技术的发展过程可以用图 4-4 表示。

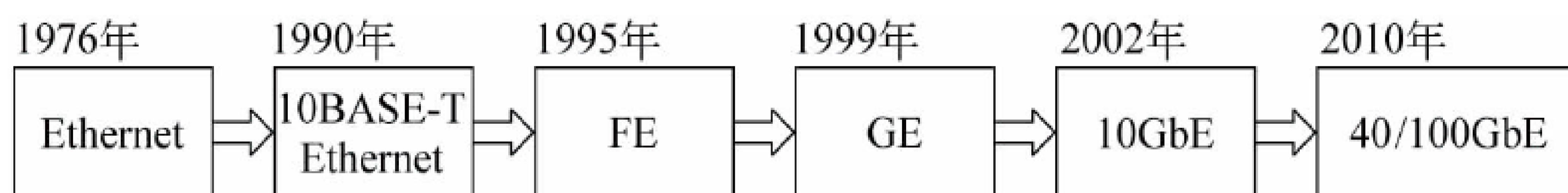


图 4-4　Ethernet 技术的发展过程

（4）将共享介质方式改为交换方式，这就导致了交换式局域网技术的发展。交换局域网的核心设备是局域网交换机，它可以在交换机多个端口之间同时建立多个并发连接。这就导致局域网被分为两类：共享式局域网（Shared LAN）和交换式局域网（Switched LAN）。

（5）将一个大型局域网划分成多个用网桥或路由器互联的小型局域网，这就导致了局域网互联技术的发展。网桥、交换机与路由器可以隔离子网之间的广播通信量。通过减少每个子网内部结点数 N 的方法，使每个局域网的网络性能得到改善。

问题 4-3：如何认识 Ethernet 的未来发展趋势？

Ethernet 已经成为办公自动化环境组网的首选技术，全世界已经有数亿台计算设备接入到 Ethernet 中。Ethernet 技术的广泛应用造就了一个巨大并且竞争激烈的市场，成为网络技术研发的一个重点领域。Ethernet 技术发展趋势可以总结为以下几点。

1. 速率更高

从 1980 年第一个速率为 10Mbps 的 Ethernet 标准出现之后的三十多年中，高速 Ethernet 沿着 100Mbps、1Gbps、10Gbps、40Gbps 到 100Gbps 的步伐一步步地前进，2013 年，IEEE 又成立了 802.3bs 工作组，研究速率达到 400Gbps 的下一代 Ethernet 标准与技术。

2. 应用更广

高速 Ethernet 与光 Ethernet、城域 Ethernet 技术的发展，使得 Ethernet 的应用从局域网逐步扩大到城域网与广域网，正在向覆盖范围越来越广的方向发展；同时从组建办公环境的局域网，向组建近距离、高吞吐量、低延时的大型高性能计算机系统、存储区域网、云计算

平台等后端计算机机房网络的方向发展;工业 Ethernet 正在广泛应用于工业自动化领域,成为工业 4.0 发展的重要支撑技术。

3. 与无线局域网兼容

无线局域网 WLAN 以微波、激光与红外等无线信道取替传统 Ethernet 中的同轴电缆、双绞线与光纤,实现移动结点的物理层与介质访问控制子层的功能。IEEE 在 802.11 无线局域网标准制定过程中,一直保持与 802.3 标准的 Ethernet 兼容,因此有人将无线局域网“Wi-Fi”称为“无线 Ethernet”。

4. 更环保

2000 年的一份研究报告指出:从 100Mbps 到 1Gbps 的 Ethernet 端口耗电约 4W。如果美国 1.6 亿台接入 Ethernet 的计算机在网络空闲时进入低功率模式,一年可以节能 2.4 亿美元的电费。2010 年 9 月,IEEE 发布的 802.3az 支持“Energy Efficient Ethernet,EEE”标准。EEE 标准通过在没有数据需要发送时关闭 Ethernet 接口的方式,达到节约能源的目的。目前 802.3az 标准已经在一些使用铜缆的 Ethernet 中使用,正在向光纤介质系统中扩展。

问题 4-4:协议 ALOHA、CSMA、CSMA/CD 与 CSMA/CA 是什么关系?

在研究生入学统考大纲中要求学生掌握随机型访问介质控制的知识点,掌握随机访问控制协议与轮询访问控制协议,在随机访问控制协议中了解 ALOHA 协议与 CSMA 协议,以及 CSMA/CD 协议与 CSMA/CA 协议。理解它们之间的关系,需要注意以下几个问题。

1. ALOHA 协议

在 ALOHA 协议、CSMA 协议,以及 CSMA/CD 协议与 CSMA/CA 协议中,ALOHA 协议是最早出现的随机型访问介质控制协议,CSMA 协议是在它的基础上发展起来的,但是由 CSMA 协议基础上延伸出来的 CSMA/CD 协议用在 Ethernet 中,CSMA/CA 协议用于无线局域网 WLAN 中。IEEE 802.3 标准是在 CSMA/CD 协议的基础上完成的;IEEE 802.11 标准是在 CSMA/CA 协议基础上完成的。Ethernet 的核心技术——随机型共享总线的介质访问控制 CSMA/CD 方法的设计思想是来源于 ALOHANET 的 ALOHA 协议。

2. ALOHA 协议

ALOHANET 出现在 20 世纪 60 年代末期。夏威夷大学的 Norman Abramson 和同事们为了在位于夏威夷各个岛屿上的不同校区之间进行计算机通信,研究了一种以无线广播方式工作的分组交换网。ALOHANET 使用的是一个共用的无线电信道,支持多个结点对一个共享的无线信道的“多路访问”。当多个终端向 IBM 360 主机传输数据时,就可能出现两个或两个以上的终端同时争用一个通信信道而产生“冲突”的情况。解决“冲突”的办法只有两种:一种是集中控制的方法,另一种是分布控制的方法。ALOHA 协议是一种分布式控制、随机争用型的系统。随机争用型访问控制方法经历了从纯 ALOHA、时间片 ALOHA 到载波侦听多路访问 CSMA 方法的演化过程。

1) 纯 ALOHA

ALOHA 系统最初是在无线共用信道上实现的,后来研究出多种类似 ALOHA 方法的系统。为了区分各种 ALOHA 方法,通常将最原始的、不对发送数据的结点做任何约束的方法称为纯 ALOHA(Pure ALOHA)。纯 ALOHA 可以工作在无线信道,也可以工作在总线型的局域网中。

2）时间片 ALOHA

纯 ALOHA 的特点是不对结点发送数据的时间做任何约束，那么它的吞吐率 S 的最大值只能达到 0.184。1972 年，Roberts 发表了可以使 ALOHA 的吞吐率 S 提高一倍的方法。改进的方法很简单：将结点可以利用的传输介质发送时间划分为等长的时间片，每个时间片用于发送一个数据帧，同时规定所有需要发送数据帧的结点只能在每个时间片开始时发送。这种方法称为时间片 ALOHA 或 SALOHA（Slotted ALOHA）。在吞吐率 S 方面，时间片 ALOHA 方法比纯 ALOHA 方法有很大的改善。

3. CSMA

由于时间片 ALOHA 只控制信道的时间片，而对结点的发送过程没有任何控制，因此信道的最大利用率只能达到 37%。在局域网中，任何一个结点在发送数据之前可以检测到是否已经有结点在使用共用传输介质。因此，人们会想到：如果对局域网中的发送过程采取一定的控制，可以提高共用传输介质的最大信道利用率，这个方法就是载波侦听多路访问（CSMA）。所谓的多路访问是指由多个结点共用一个共享的传输介质，载波侦听是指各个结点在发送数据之前需要确认是否已有结点使用共用信道传输数据，通过主动避免冲突的方法去进一步提高信道利用率。1975 年，Kleinrock 与 Tobagi 对多种 CSMA 方法进行了分析。

由于每个结点都能在发送之前，会监听其他结点是否在发送数据帧，如果已经有结点正在发送数据帧，那么这个结点就暂时不发送数据，从而减少发生冲突的可能性。根据采用的具体方法上的差异，CSMA 方法又可以分为以下两种类型。

1）非坚持 CSMA

当一个结点准备发送数据帧时，该结点开始监听信道。如果监听到信道忙，已经有其他结点在利用共用信道发送数据，就不再坚持监听状态，而是等待一个随机延迟的时间后，再重复上述过程。如果通过载波侦听发现信道空闲，则发送数据帧。

2）1-坚持 CSMA 和 P-坚持 CSMA

坚持 CSMA 的特点是：当监听到信道忙时，仍然坚持监听下去，一直坚持听到信道空闲为止。当信道空闲后，可以采取两种不同的策略。第一种方法是发现信道空闲就立即发送数据帧，这种方法的出发点是抓紧一切有利时机发送数据。但是，如果有两个或两个以上结点同时监听信道，则一旦信道空闲就必然有多个结点同时发送数据帧而产生冲突。为了避免这种情况，可以采用第二种折中的方法，那就是当结点听到信道空闲时，就以概率 P 来发送数据帧，而以概率 $1-P$ 延迟一个时间 τ 来重新监视听信道，这种方法被称为 P-坚持 CSMA。第一种方法被称为 1-坚持 CSMA，因为这种方法在结点监听到信道空闲时，以概率 $P=1$ 来发送数据帧。

4. CSMA/CA

对 ALOHA 方法的改进，一种方法是从发送结点如何决定数据帧是否发送入手。CSMA 方法要求每个结点在发送之前监听共享传输介质，以确定是否已经有其他结点在发送数据帧。如果已经有结点正在发送数据帧，这个结点就暂时不发送数据，从而减少发生冲突的可能性，提高了共享传输介质的信道利用率。这种方法叫作载波侦听多路访问/冲突避免（CSMA/CA），IEEE 802.11 无线局域网的 MAC 层就采用这种方法。

5. CSMA/CD

对 ALOHA 改进的第二种方法是：如果有两个结点都侦听到共享传输介质空闲，同时发送了数据帧，并且它们几乎同时检测到冲突发生。第二点改进是结点一旦检测到冲突，不是继续发送完数据帧，而是尽快停止冲突帧的传送，这种方法就是带有冲突检测的载波侦听多路访问(CSMA/CD)方法。这是一种分布式共享介质控制方法。IEEE 802.3 的 Ethernet 局域网的 MAC 层就采用这种方法。

问题 4-5：CSMA/CD、Token Bus 与 Token Ring 有哪些共同之处?

学习局域网技术必然涉及对三种主要局域网类型的比较。这三种局域网类型是：符合 IEEE 802.3 标准的 Ethernet、符合 IEEE 802.4 标准的令牌总线网(Token Bus)与符合 IEEE 802.5 标准的令牌环网(Token Ring)。

(1) 对于这三种类型的局域网，它们共同遵循 IEEE 802 体系结构，也就是说都是按物理层、介质访问控制子层与逻辑链路控制子层的层次结构模型来设计。它们的物理层、介质访问控制子层不同，而在逻辑链路控制子层采取相同的协议标准。

(2) 它们在数据传输技术上都采用以广播方式共享的共用传输介质。传输介质类型主要有双绞线、同轴电缆与光纤。

(3) 决定它们性能的三个因素同样是：拓扑、传输介质与介质访问控制方法。

(4) 影响局域网性能与应用领域的主要因素是所采用的介质访问控制方法。三种局域网都采用了分布式访问控制方法，在网络中没有集中控制结点。

(5) 介质访问控制方法是指控制多个结点利用共用传输介质发送和接收数据的方法，它是所有“共享介质”类型的局域网都必须解决的共性问题。介质访问控制方法要解决以下几个问题：哪个结点发送数据? 发送时会不会出现冲突? 出现冲突时怎么办?

问题 4-6：如何比较 CSMA/CD、Token Bus 与 Token Ring 的性能?

理解这个问题需要注意以下几点。

(1) 比较 CSMA/CD、Token Bus 与 Token Ring 的性能可以用理论分析的方法，并可以通过实验环境的实测来进行比较。

(2) 图 4-5 给出了 Token Bus、Token Ring 与 CSMA/CD 方法在不同网络通信负荷情况下实际能达到的数据传输速率的比较。图 4-5 的横坐标与纵坐标的单位均为 Mbps。曲线的测试条件是：三种局域网都传输长度为 2000b 的帧，网络中有 100 个结点，并且 100 个结点都在利用共享的传输介质发送数据帧。图中的横坐标表示网络的数据传输速率(单位是 Mbps)，它实际上表示的是不同的物理层数据帧的发送时钟。因为如果物理层发送一个比特的时间是 0.1μs，那么它对应的数据传输速率为 10Mbps。图中的纵坐标表示网络实际能成功发送数据帧的最大数据传输速率(单位是 Mbps)。由于发送时存在冲突而必须采用介质访问控制方法，因此网络实际能成功发送数据帧的最大数据传输速

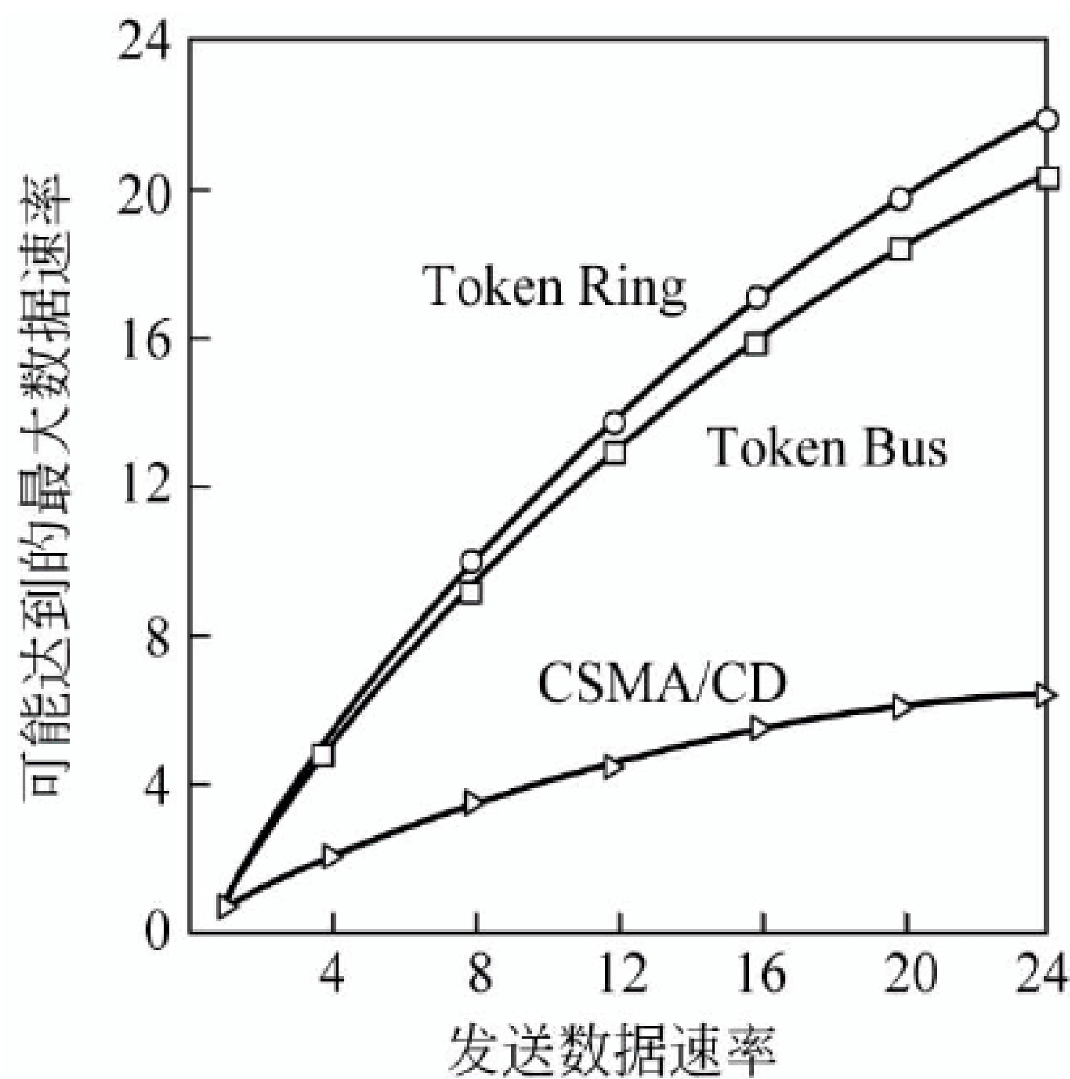


图 4-5 不同通信负荷的实际数据传输速率的比较

率,反映了不同的介质访问控制方法对信道利用率的影响。我们可以比较在数据传输速率为10Mbps时,使用CSMA/CD方法的网络最大可以达到的数据传输速率大约为4.47Mbps;使用Token Bus方法的网络最大可以达到的数据传输速率大约为8.94Mbps;使用Token Ring方法的网络最大可以达到的数据传输速率大约为9.41Mbps。

(3) 从对图4-5的分析中,可以得出以下几个重要的结论。

① 在使用CSMA/CD方法的Ethernet中,假设有100个结点,每个结点都以10Mbps的数据传输速率发送长度为1000b的帧,可能由于碰撞而造成帧的发送失败,实际能利用共享传输介质成功发送的数据最多为4.47Mbps。也就是说,在这样的网络负载的条件下,对于数据传输速率为10Mbps的Ethernet来说,能够成功发送数据的实际速率最多可达到4.47Mbps。

② Token Bus与Token Ring在网络通信负荷较重时,表现出很好的吞吐率与较低的传输延迟。而CSMA/CD在网络通信负荷增大时,由于冲突增多而造成网络吞吐率下降和传输延迟增加。因此,使用Token Bus与Token Ring方法的网络适用于通信负荷较重的应用环境,而使用CSMA/CD方法的网络一般用于通信负荷较轻、实时性要求不高的应用环境。

③ Token Bus与Token Ring在网络通信重负荷中表现很好,是以复杂的环控制功能为代价来实现的。要完成复杂的环控制功能,Token Bus与Token Ring的网卡与联网设备比较复杂,硬件造价高,组网的费用远远超过使用CSMA/CD方法的Ethernet。

④ 随着个人计算机的广泛应用,办公自动化环境中计算机联网的需求快速增长,组网费用低廉的Ethernet正好能适应这种对传输延迟要求不高的应用,因此Ethernet技术相对于其他两种环网技术有很大优势。到20世纪90年代,局域网市场激烈竞争的局面已经明朗,Ethernet产品基本上垄断了市场,Ethernet几乎成了局域网的代名词。但是,需要注意的是,使用CSMA/CD方法的Ethernet只适用于对传输实时性要求不高的应用,例如机关、企业办公自动化和学校的学习环境。

⑤ 对于工业环境中,例如机器人控制、制造业设备与仪表的现场控制,这类对数据传输实时性要求严格的应用,一般使用基于Token Bus与Token Ring原理设计的局域网,而不能使用Ethernet。不同特性的网络应该有相应的应用领域,不可能出现能适应各种应用需求的技术。

问题4-7:Token Bus、Token Ring与CSMA/CD具有哪些特点?

1. CSMA/CD方法的特点

与确定型介质访问控制方法Token Bus、Token Ring比较,CSMA/CD方法具有以下几个主要的特点。

(1) CSMA/CD介质访问控制方法算法简单,并且易于实现。目前,有多种VLSI可以实现CSMA/CD方法,这样有利于降低Ethernet成本和扩大应用范围。

(2) CSMA/CD是一种用户访问总线时间不确定的随机竞争总线的方法,适用于办公自动化等对数据传输实时性要求不严格的应用环境。

(3) CSMA/CD在网络通信负荷较低时表现出较好的吞吐率与延迟特性。但是,当网络通信负荷增大时,由于冲突增多,网络吞吐率下降、传输延迟增加,因此CSMA/CD方法一般用于通信负荷较轻的应用环境中。

2. Token Bus、Token Ring 的特点

与 CSMA/CD 比较,Token Bus、Token Ring 的特点主要表现在以下几个方面。

(1) Token Bus 或 Token Ring 网中结点两次获得令牌之间的最大间隔时间确定,因此适用于对数据传输实时性要求较高的应用环境(例如生产过程控制)。

(2) Token Bus 与 Token Ring 在网络通信负荷较重时,表现出很好的吞吐率与较低的传输延迟,因此适用于通信负荷较重的应用环境。

(3) Token Bus 与 Token Ring 的不足之处是都需要复杂的环维护功能,实现起来比较困难。

问题 4-8:支持 TCP/IP 的 IEEE 802 局域网和城域网都包括哪些协议标准?

需要注意的是:IEEE 802 委员会为了制定局域网标准而成立了一系列组织,例如,制定某类协议的工作组(WG)或技术行动组(TAG)。它们研究和制定的标准统称为 IEEE 802 标准。随着局域网技术的发展,IEEE 802.4WG、IEEE 802.6WG、IEEE 802.7WG、IEEE 802.12WG 等工作组已停止工作。目前,最活跃的工作组是 IEEE 802.3WG、IEEE 802.11WG、IEEE 802.15WG 等。

IEEE 802 委员会公布了很多标准,这些协议可以分为三类:定义了局域网体系结构、网络互联,以及网络管理与性能测试的 IEEE 802.1 标准;定义了逻辑链路控制 LLC 子层功能与服务的 IEEE 802.2 标准;定义了不同介质访问控制技术的相关标准。

不同介质访问控制技术的相关标准曾经多达 16 个。随着局域网技术的发展,一些过渡性技术在面对市场的检验中逐步被淘汰或很少应用,目前应用最多和正在发展的标准主要有四个,其中三个是无线局域网的标准。4 个主要的介质访问控制协议标准如下。

(1) IEEE 802.3 标准:定义 CSMA/CD 总线介质访问控制子层与物理层标准。

(2) IEEE 802.11 标准:定义无线局域网访问控制子层与物理层标准。

(3) IEEE 802.15 标准:定义近距离个人区域无线网络访问控制子层与物理层标准。

(4) IEEE 802.16 标准:定义宽带无线局域网访问控制子层与物理层标准。

目前,IEEE 802.2 标准关于逻辑链路控制 LLC 子层的协议已经很少被使用,网络层 IP 分组直接封装到 IEEE 802.3、IEEE 802.11 等协议帧的数据字段之中。图 4-6 给出了一个简化的 IEEE 802 协议结构。

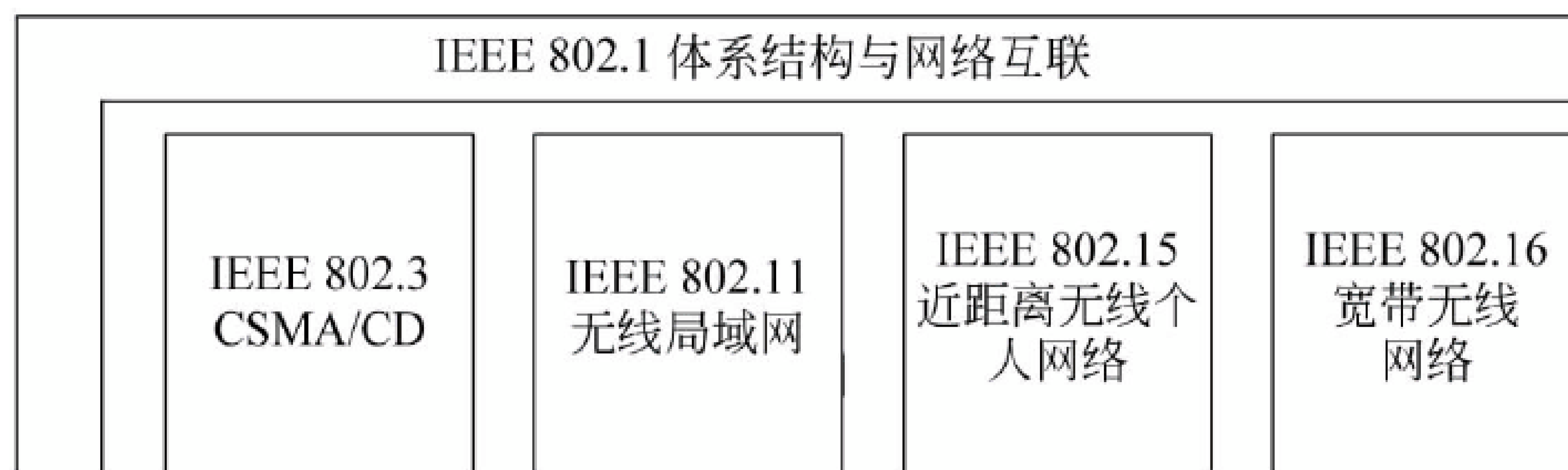


图 4-6 IEEE 802 协议结构

当前最流行的局域网与城域网标准包括 IEEE 802.3(Ethernet)和 IEEE 802.11(Wi-Fi)。这些标准随着时间推移而演变,经过修订名称也发生改变,并最终被纳入修订过的标准。表 4-1 给出了支持 TCP/IP 的相关 IEEE 802 局域网和城域网标准的比较完整的列表。

表 4-1　有关 TCP/IP 的局域网和城域网 IEEE 标准(2011 年)

名　　称	描　　述	参考文献
802.1ak	多注册协议(MRP)	802.1AK—2007
802.1AE	MAC 安全(MACSec)	802.1AE—2006
802.1AX	链路聚合(以前的 802.3ad)	802.1AX—2008
802.1d	MAC 网桥	802.1D—2004
802.1p	流量类/优先级/QoS	802.1D—2004
802.1q	虚拟网桥的局域网/MRP 的更正	802.1Q—2005/Corl—2008
802.1s	多生成树协议(MSTP)	802.1Q—2005
802.1w	快速生成树协议(RSTP)	802.1D—2004
802.1X	基于端口的网络访问控制(PNAC)	802.1X—2010
802.2	逻辑链路控制(LLC)	802.2—1998
802.3	基本以太网和 10Mbps 以太网	802.3—2008
802.3u	100Mbps 以太网("快速以太网")	802.3—2008
802.3x	全双工操作和流量控制	802.3—2008
802.3z/802.3ab	1000Mbps 以太网("千兆以太网")	802.3—2008
802.3ae	10Gbps 以太网("十千兆以太网")	802.3—2008
802.3ad	链路聚合	802.1AX—2008
802.3af	以太网供电(15.4W)	802.3—2008
802.3ah	以太网接入	802.3—2008
802.3as	帧格式扩展(2000B)	802.3—2008
802.3at	以太网供电增强(PoE+,30W)	802.3at—2009
802.3ba	40/100Gbps 以太网	802.3ba—2010
802.11a	在 5GHz 的 54Mbps 的无线局域网	802.11—2007
802.11b	在 2.4GHz 的 11Mbps 的无线局域网	802.11—2007
802.11e	针对 802.11 的 QoS 增强	802.11—2007
802.11g	在 2.4GHz 的 54Mbps 的无线局域网	802.11—2007
802.11h	频谱/电源管理扩展	802.11—2007
802.11i	安全增强/代替 WEP	802.11—2007
802.11j	日本的 4.9～5.0GHz 操作	802.11—2007
802.11n	在 2.4 与 5GHz 的 6.5～600Mbps 的无线局域网,使用可选的 MIMO 和 40MHz 信道	802.11n—2009
802.11s(草案)	网状网,拥塞控制	开发中
802.11y	在 3.7GHz 的 54Mbps 的无线局域网(许可的)	802.11y—2008

续表

名　　称	描　　述	参考文献
802.16	微波存取全球互通技术(WiMax)	802.16—2009
802.16d	固定的无线城域网标准(WiMax)	802.16—2009
802.16e	固定/移动的无线城域网标准(WiMax)	802.16—2009
802.16h	改进的共存机制	802.16h—2010
802.16j	802.16 中的多跳中继	802.16j—2009
802.16k	802.16 网桥	802.16k—2007
802.21	介质独立切换	802.21—2008

问题 4-9：术语辨析：介质、介质访问与介质访问控制。

(1) 在局域网的讨论中，经常会使用术语“介质(Media)”。在这里，“介质”指的是“传输介质”，如同轴电缆、双绞线、光纤、无线通信信道，而不是通常意义上的 RAM、ROM、磁盘与光盘等存储介质。

(2) 术语“介质访问”是指：连接在共享传输介质上的计算机，利用传输介质发送与接收数据的过程。

(3) 术语“介质访问控制”是指：如何使得连接在共享传输介质上的多台计算机，在利用共享的传输介质发送数据时不出现冲突，以及出现冲突之后怎么解决的分布式控制方法。

问题 4-10：为什么要将 Ethernet 的基本工作原理与实现技术作为重点内容系统讨论？

主教材将 Ethernet 的基本工作原理与实现技术作为本章的重点内容进行系统讨论的理由主要有以下几点。

(1) Ethernet 在与 Token Bus、Token Ring 的竞争中脱颖而出，成为局域网领域的主流技术。所有的操作系统都内置了各种类型 Ethernet 网卡的驱动程序。各种网络管理与测试设备都是以 Ethernet 为基本联网环境。Ethernet 物理层 10BASE-T 标准奠定了结构化布线系统的基础。

(2) Ethernet 技术越来越重要。我们在办公环境中需要将计算机接入到 Ethernet 中，家庭计算机需要通过 Ethernet 网卡的 100BASE-T 接口连接到 ADSL 调制解调器或无线路由器，进而接入 ISP。

(3) Ethernet 传输速率从传统的 10Mbps，发展到 100Mbps、1Gbps、10Gbps、40Gbps、100Gbps，光纤的应用使得 Ethernet 覆盖范围从局域网逐步应用到城域网、广域网。无线局域网 802.11 标准的帧结构设计考虑与 802.3 帧的兼容性问题。Ethernet 已经成为未来计算机网络技术发展与应用的基础。

(4) CSMA/CD 已经成为局域网 MAC 层分布式控制最经典的算法。但是很多教科书对于 CSMA/CD 算法分析不深入，学生学习之后一般只能接受一些基本的说法和部分术语，对于怎么实现 CSMA/CD 算法，如何从计算机外设的角度去说明 Ethernet 的网卡设计与实现技术，如何编写网卡驱动程序，以及如何在 MAC 之上进行网络软件编程会感到茫然。

（5）主教材采取从 Ethernet 研究发展的角度，系统地讨论了 Ethernet 技术从传统的10Mbps，发展到100Mbps、1Gbps、10Gbps、40Gbps、100Gbps 的发展过程，尤其是对从计算机外设的涉及与实现方法角度，去说明 Ethernet 网卡的设计与实现技术，为学生系统解析一种主流局域网技术的研究、设计与实现的方法，促进理论与实际的结合。通过网络硬件实验与网络软件编程的训练，锻炼和提高学生实际工作能力。

因此，网络课程在底层以 Ethernet 技术为突破口，切入到网络实现技术的角度讨论问题，用学生具备的电子学知识去解释复杂的实现技术，可以破除学生对网络技术的神秘感，启发学生的学习兴趣。

问题 4-11：对 Ethernet 的 CSMA/CD 工作原理的描述有几种方法？

不同的教科书对 Ethernet 的 CSMA/CD 工作原理的描述有两种基本的方法：状态图法与流程图法。

1. CSMA/CD 状态图

为了有效地实现分布式总线访问控制策略，CSMA/CD 设计的基本原则为：凡是不发送帧的结点都处于接收状态；发送结点执行"先听后发，边听边发，冲突停止，延迟重发"的 4 步流程。图 4-7 给出了 CSMA/CD 的结点工作状态图。状态图很好地表达出 Ethernet 的 MAC 层工作原理。

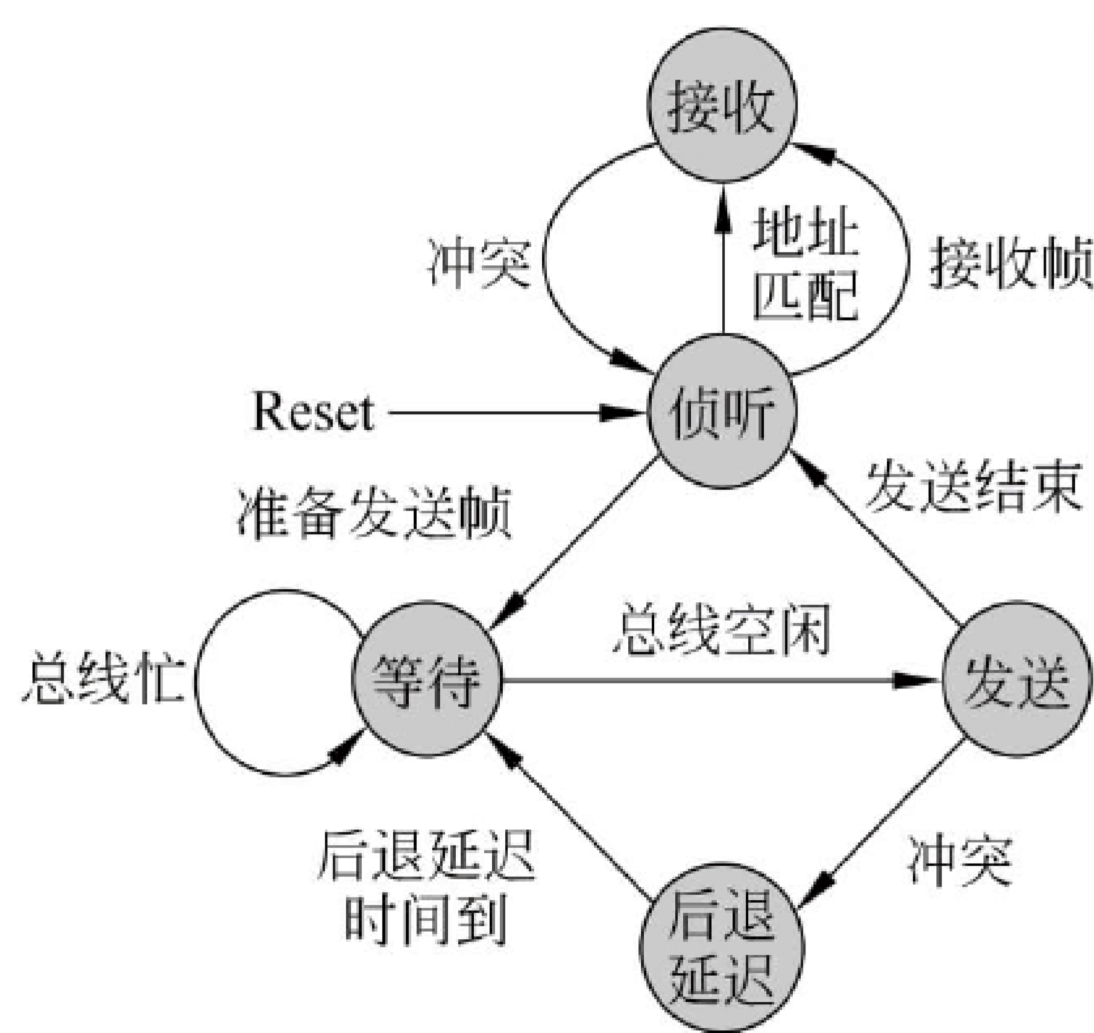


图 4-7　CSMA/CD 的结点工作状态图

从状态图 4-7 中可以看出，侦听状态是一个初始状态。任何一个接入 Ethernet 中的结点，只要不发送数据，就处于侦听状态。结点接收帧的过程是被动的，只要总线上有发送帧出现，它就要接收。当结点成功地接收到一个帧之后，根据接收帧的目的地址判断是不是本结点的 MAC 地址，或者本组地址及广播地址来确定地址是否匹配。如果地址匹配，则存储接收帧，并通知高层处理该帧。如果地址不匹配，则丢弃接收帧，重新回到接收状态。

如果一个结点准备发送帧，那么它首先要确定总线是不是忙。如果总线忙，则处于等待状态；如果总线不忙，则进入发送状态。当结点发送一个帧之后，结点需要判断发送成功还是失败。如果发送成功，并且没有新的帧需要发送，它将转回到侦听状态。如果出现冲突，则进入后退延迟状态。当后退延迟时间到达时，则重新进入等待状态，循环执行上述过程，直至发送成功，或冲突过多而失败。

2. CSMA/CD 发送流程与接收流程图

图 4-8 给出了 CSMA/CD 的结点工作流程图。主教材采用的是 CSMA/CD 的结点工作流程图。使用 CSMA/CD 的结点工作流程图对于网络编程有指导意义。

问题 4-12：Ethernet 是通过什么方法来判断总线的忙闲状态与冲突的？

了解这些知识对于理解 Ethernet 基本工作原理，了解 Ethernet 为什么能够广泛应用很有帮助。

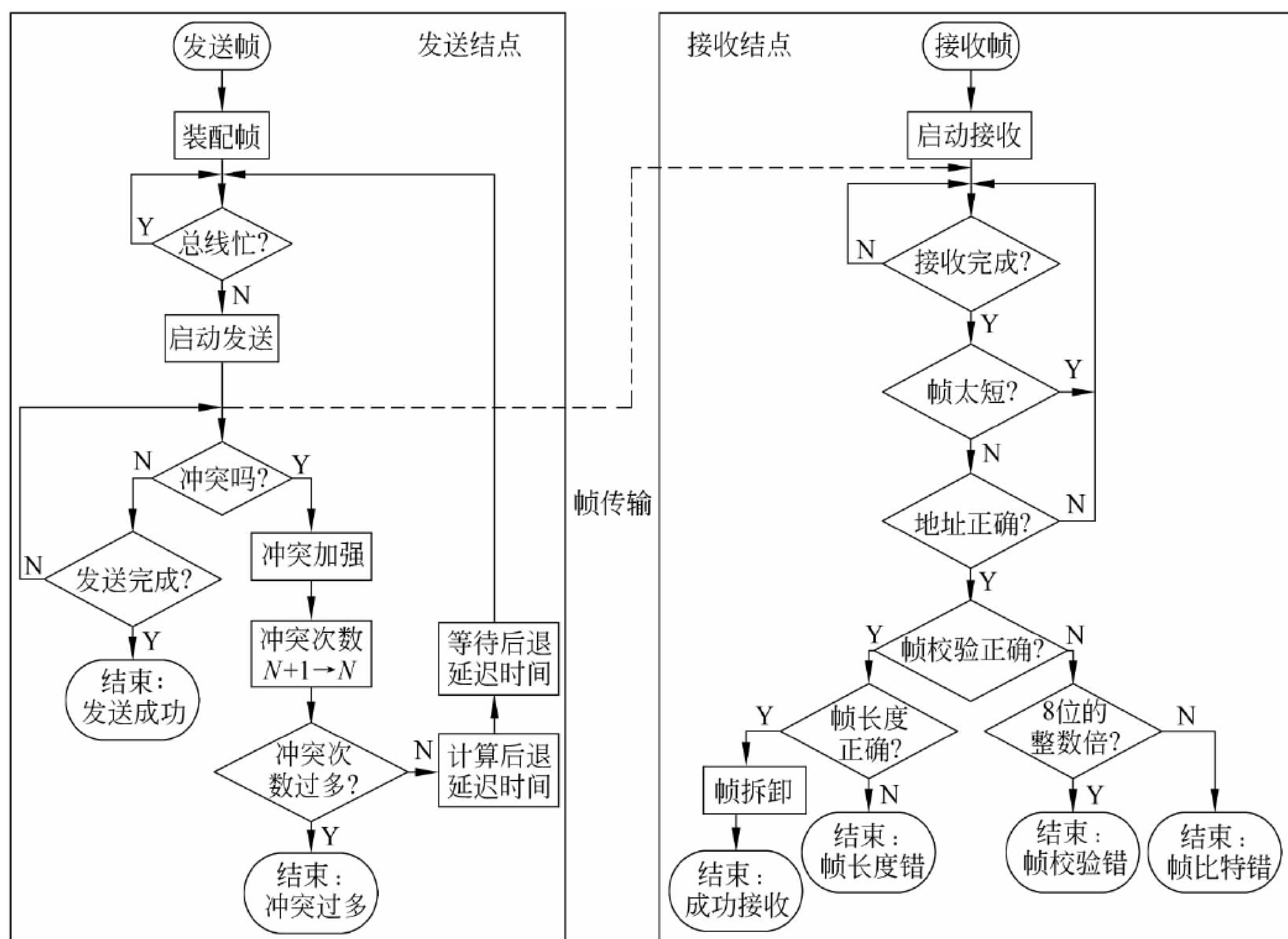


图 4-8　发送流程与接收流程图

1. Ethernet 网卡判断总线忙闲状态的方法

Ethernet 网卡判断总线忙闲状态的方法很简单。网卡只要监听总线,总线上有信号发送时,它的解码电路总有时钟脉冲信号输出。一旦总线没有信号发送时,它的电平保持不变,接收电路的解码时钟就没有时钟脉冲信号输出,这时就可以判断总线空闲。图 4-9 给出了总线忙闲与解码时钟的关系示意图。

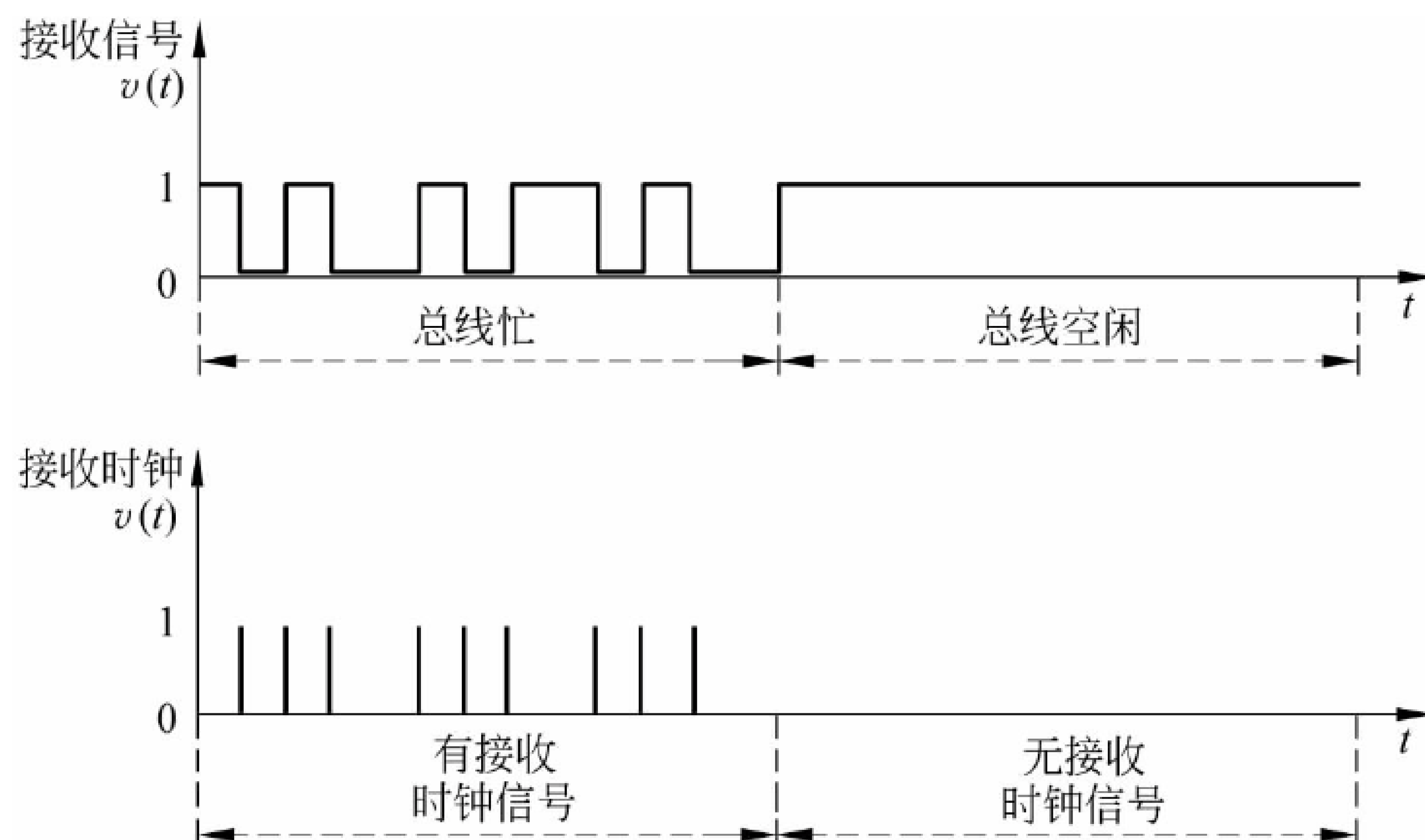


图 4-9　总线忙闲与解码时钟的关系

2. Ethernet 判断冲突的方法

Ethernet 判断是否出现冲突的方法有两种:一是比较法;二是编码为例法。这两种方

法都要求发送结点在发送数据帧的同时，接收总线上出现的信号。

比较法的思路是：如果接收到的比特序列与自己发送的比特序列不一致，说明总线上还有其他的结点在发送信号，因此可以判断出现冲突，此次发送失败。如果一致，表明只有自己在发送数据，不存在冲突。

编码为例法的理论依据是：如果两路或两路以上的 Manchester 编码信号同时出现在总线上，它们叠加的结果，其波形一定不符合 Manchester 编码规律。

因此，这两种方法在电子线路上很容易实现。因此，说 Ethernet 的 CSMA/CD 算法实现简单是有道理的。

问题 4-13：在什么情况下 CSMA/CD 算法是正确的？

这个问题换一种提法是：推导 CSMA/CD 有没有前提？其实，这是深入理解 CSMA/CD 工作原理与实现技术最核心的问题。这里涉及“冲突”与“冲突窗口”的概念。很多关于物理层协议的参数与组网的限制条件，以及帧结构的设计都是从这一点出发的。

理解 CSMA/CD 算法：载波侦听、冲突检测、冲突处理、随机延迟重发的基本工作原理，需要注意以下几个问题。

(1) 每个结点要利用总线发送数据时，首先要通过“载波侦听”来确定总线是否空闲。载波侦听方法并不能完全消除冲突。

(2) “冲突检测”引出了冲突窗口的概念。冲突窗口是 Ethernet 组网技术的一个重要参数。冲突窗口是指一个结点在开始发送数据后，可以检测到冲突的最长时间。换句话说，如果一个结点在发送信号后，在超过冲突窗口固定的时间后仍没有检测出冲突，就说明此次发送成功。为了保证任何一台主机在发送一帧的过程中都能够检测到冲突，就要求发送一个最短帧的时间都要超过冲突窗口的时间。

(3) 冲突窗口与最小帧长度之间的关系。

如果最短帧长度为 L_{min}，主机发送速率为 S，发送短帧所需要的时间为 L_{min}/S。冲突窗口值为 $2D/V$。要求发送一个最短帧的时间都要超过冲突窗口的时间，即

$$L_{min}/S \geqslant 2D/V$$

那么，总线长度与最小帧长度、发送速率之间的关系为

$$D \leqslant VL_{min}/2S$$

可以根据总线长度、发送速率与电磁波传播速度，估算出最小帧长度。

IEEE 802.3 协议将 10Mbps 的冲突窗口定为 51.2μs。

(4) 如果在发送数据的过程中检测出冲突，为了解决信道争用冲突，发送结点要进入“停止发送”“随机延迟后重发”的流程。随机延迟后重发流程的第一步是发送“冲突加强信号”。发送冲突加强信号的目的是确保有足够的冲突持续时间，使得网中所有结点都能检测出冲突并立即丢弃冲突帧，减少因冲突浪费的时间和提高信道利用率。

(5) 我们注意这句话：“如果一个结点在发送信号后，在超过冲突窗口固定的时间后仍没有检测出冲突，就说明此次发送成功。”这句话换一种说法是：如果一个结点发送帧需要的时间短于冲突窗口，那么这个结点不可能知道它所发送的帧是否出现冲突，也就不能够确定这一帧发送是否成功。基于这点，IEEE 802.3 协议将 10Mbps 的冲突窗口定为 51.2μs，将帧的最短长度定为 64B。如果帧长度短于 64B，就需要添加填充字节，其目的就是要在一帧发送完之前必须检测到是否出现冲突。

(6) 在早期使用粗同轴电缆的10BASE-5标准组建一个Ethernet时的“4-4-3规则”,即限制局域网中两个通信的结点之间的路径,最多通过5个长度为500m的粗同轴电缆段,最多经过4个中继器,5个缆段中只有3个连接有结点。做出这个限制的根据是:如果两个结点之间的路径和大于以上条件时,信号传输延迟将大于冲突窗口,这样CSMA/CD算法的前提条件“在一帧发送的过程中,结点能够及时检测到冲突”的条件被破坏,因此“4-4-3规则”被作为用10BASE-5标准的粗同轴电缆组建一个Ethernet的设计依据的原因。

通过4个中继器连接的5个500m的电缆段最大长度为2500m。电磁波在铜缆中传播速度约为0.77C(C为光速),可得到64B采用10Mbps时的传输时间为$64\times8/10\,000\,000=51.2\mu s$。Ethernet设计者确定最小帧长度基于:所有结点在发送帧的最后一位传输时都能够检测到是否发生冲突。

由于10BASE-5、10BASE-2标准目前已经基本上不使用,因此在主教材中没有涉及这一部分内容。但是,为了理解CSMA/CD算法应该了解这些知识。

(7) 需要注意的是:在高速Ethernet中,如果速率提高10倍(100Mbps),那么同样发送64B所需的时间就缩短到十分之一,因此高速Ethernet中需要保持与传统10Mbps的Ethernet兼容,就必须采取相应的解决办法。这个问题在高速Ethernet的讨论中会涉及。采用交换式Ethernet技术就不需要采用CSMA/CD控制方法。

问题4-14:如何理解CSMA/CD执行过程中的随机后退延迟算法?

理解CSMA/CD中的随机后退延迟算法,需要注意以下几个问题。

(1) Ethernet协议规定当发生冲突时,发送结点由随机后退延迟算法来决定后退延迟,再次进入总线忙侦听的延迟时间。算法的全称为:截止二进制指数后退延迟算法。

(2) 后退延迟算法是截止二进制指数后退延迟算法。该算法可以表示为

$$\tau = 2^k \times R \times a$$

其中,τ为结点重新发送需要的后退延迟时间,a为冲突窗口值,R为随机数。如果一个结点需要计算后退延迟时间,则需要以其地址为初始值产生一个随机数R。冲突窗口a值是$51.2\mu s$。从计算公式中可以看出,结点重发后退的延迟时间是冲突窗口值的整数倍,并与以冲突次数为二进制指数的幂值成正比。

(3) “截止二进制指数”后退延迟算法的特点体现在限定了作为二进制指数k的范围。

它定义了$k=\min(n,10)$。如果重发次数$n<10$,则取$k=n$;如果重发次数$n\geqslant10$时,则k取值为10。例如,如果第一次冲突发生,则重发次数$n=1$;由于$n<10$,则取$k=1$,即在冲突后两个时间片后重发。如果第二次冲突发生,则重发次数$n=2$,由于$n<10$,则取$k=2$,即在冲突后4个时间片后重发。在$n<10$时,随着n的增加,重发延迟时间按$2n$幂值增长。在$n>10$时,重发延迟时间不再增长。由于限制了二进制的指数k的范围,那么第n次重发延迟是分布在0与$2^{\min(n,10)}-1$个时间片之间,最大可能延迟时间为1023个时间片,每个时间片按冲突窗口值为$51.2\mu s$计算,后退延迟时间为52.4ms(假设随机数的值为1)。也就是说,当冲突出现10次时,需要延迟52.4ms,那么出现第11次冲突时,也只要延迟52.4ms。在后退延迟时间到达后,结点将重新判断总线忙、闲状态,重复发送流程。

(4) 规定一个帧的最大重发次数为16。如果重发次数超过16,则认为线路故障,系统进入“冲突过多”结束状态。如果重发次数$n\leqslant16$,则允许结点随机延迟再重发。如果重发次数达到16,则认为线路故障,系统进入“冲突过多”结束状态。

(5) 从随机后退延迟算法可以看出,CSMA/CD 是一种随机、分布式的 MAC 控制方法。

问题 4-15:Ethernet V2.0 规范与 IEEE 802.3 标准在帧结构上有哪些差别?

理解这个问题需要注意以下几点。

(1) Ethernet V2.0 规范与 IEEE 802.3 标准在帧结构上的区别主表现在对 2B 的类型字段上。图 4-10 给出了 Ethernet V2.0 规范的帧结构示意图。Ethernet V2.0 规范是在 DEC、Intel 与 Xerox 公司合作研究的 Ethernet 协议的基础上改进而成,因此有些文献中将 Ethernet V2.0 帧结构称为 DIX 帧结构。



图 4-10　Ethernet 帧结构

(2) Ethernet V2.0 规范和 IEEE 802.3 标准中的 Ethernet 帧结构有些差别。这是因为 IEEE 802.3 标准制定时要考虑 IEEE 802.4、IEEE 802.5 等标准的兼容问题,因此"类型字段"被规定为"类型/长度字段"。在处理 IEEE 802.3 的 Ethernet 帧时,要根据"类型/长度"字段值来确定它是"类型"还是"长度"。目前,IEEE 802.4、IEEE 802.5 标准已很少使用,基本都采用 Ethernet V2.0 规范规定的 Ethernet 帧结构。在没有专门说明的情况下,本书讨论的是 Ethernet V2.0 的帧。

(3) Ethernet V2.0 标准中的 Ethernet 帧结构只有一个 2B 的"类型字段",而没有"长度字段",但是它限定了帧的最小长度。如果发送端的网络层向 MAC 子层发送的 IP 分组长度为 40B,短于帧的最小数据字段长度 48B 的规定,那么 MAC 子层在组帧之前要填充 8B,达到最小数据字段长度 48B 的要求。在组帧的过程中,加上帧头和帧尾的 18B,发送的是长度为 64B 的帧。当这个帧传送到接收端的 MAC 子层时,由于帧头没有设置长度字段,接收端的 MAC 子层并不知道发送端是否对数据字段做了填充,以及填充了多少个字节。

(4) 解决 Ethernet V2.0 规范和 IEEE 802.3 标准中的 Ethernet 帧结构的差异注意是在短帧填充字节的处理上。IEEE 802.3 标准最初将对应 2B 的"类型字段"定义为"长度字段"。由于长度字段的值不包括填充字节数,因此定义了"长度字段"可以很好地解决短帧的填充字节处理问题。但是,由于 Ethernet V2.0 标准已经广泛应用,所以 IEEE 802.3 标准在之后的修订中拿出了一个折中的方案,将 2B 定义的"长度字段"改为"长度/协议字段"。同时表示长度或协议是不矛盾的。由于 Ethernet 帧的最大长度小于 1518B,如果用十六进制表示,长度字段值一定小于 0x0600。IEEE 定义的协议字段值最小为 0x0800(IP 协议)。这样,Ethernet 的 MAC 层可以根据需要发送两种帧,可以表示上层协议的类型,也可以表示帧长度。接收端 MAC 层可根据该字段的值来解释该字段表示的意义。这样做就很好地解决了 IEEE 802.3 标准与 Ethernet V2.0 标准之间存在的差异性。

问题 4-16:为什么在计算题中 Ethernet 帧最大长度有时用 1518B,有时用 1526B?

从 Ethernet 帧结构的讨论中,可以看出 Ethernet 帧由三部分组成。

(1) 第一部分:前导码与帧前定界符,长度为 8B。

前导码是"10101010",后面是 56 位的"10101010…10101010"比特序列,共同构成了 8B

的前导码与帧前定界符，其作用是用于接收电路通过提取曼彻斯特编码的自含时钟，实现收发双方的比特同步。发送时需要有这8B的前导码与帧前定界符部分，接收后不需要保留，也不计入帧头的长度中。

(2) 第二部分：帧头部分，长度为18B。

帧头部分长度为18B(6B的目的地址、6B的源地址字段、2B的长度字段、4B的帧校验字段)。这一部分在发送和接收过程中是不变的。

(3) 第三部分：数据字段，长度在46～1500B。

最小数据字段长度为46B，不够时需要填充。最大数据字段长度为1500B。

因此，在讨论MAC层帧长度时给出的是64～1518B。这是由于接收端接收并存储的是帧头与数据字段这两部分。如果在发送端，例如在计算发送帧的时间时，就需要加上前导码与帧前定界符的8B。因此，可能引起有同一个Ethernet帧长度两种数据的情况。

问题4-17：为什么将Ethernet的MAC地址叫作物理地址？

Ethernet、Token Bus与Token Ring在MAC层标识网卡都需要使用地址，在帧的目的地址、源地址字段都需要填入地址。这个地址就叫作MAC地址。理解为什么将MAC地址又叫作Ethernet物理地址，以及物理地址的分配与管理方法，需要注意以下几个问题。

1. IEEE RAC对Ethernet物理地址的管理方法

48位的地址称为EUI-48。EUI(Extended Unique Identifier)表示扩展的唯一标识符。按照48位的Ethernet物理地址的编码方法，允许分配的Ethernet物理地址应该为2^{47}个，这个数量可以保证全球任何一个Ethernet物理地址都是唯一的。为了统一管理Ethernet的物理地址，保证每块Ethernet网卡的地址唯一，不会出现重复。IEEE注册管理委员会(Registration Authority Committee，RAC)为每个网卡生产商分配Ethernet物理地址的前三个字节，即公司标识，也称为机构唯一标识符(Organizationally Unique Identifier，OUI)。后面三个字节由网卡的厂商自行分配。从地址为ftp. ieee. org的ieee/info/info. stds. odi文件可以查到所有分配给厂商公司标识OUI的资料。

2. Ethernet物理地址的表示方法

当网卡生产商获得一个前三个字节地址分配权后，它可以生产的网卡数量是2^{24}(16 777 216)块。例如，IEEE分配给某个公司的Ethernet物理地址前个三字节可能有多个，其中一个是020100。标准的表示方法为02-01-00，在两个十六进制数之间用一个连字符隔开；该公司可以给它生产的每块Ethernet网卡分配一个后个三字节地址值，假如为2A-10-C3。这块Ethernet网卡的物理地址应该是02-01-00-2A-10-C3。图4-11给出了Ethernet物理地址的十六进制与二进制的表示方法示意图。

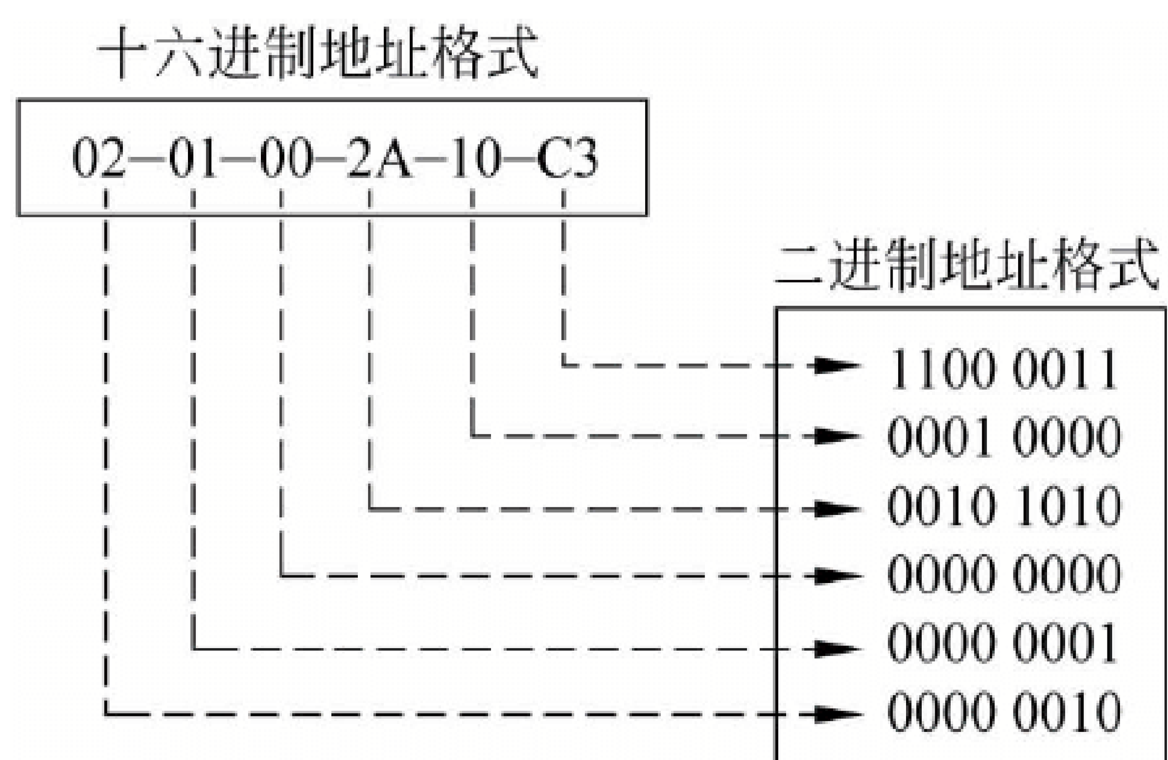


图4-11 Ethernet物理地址的十六进制与二进制表示方法

3. Ethernet 物理地址的唯一性

在网卡生产过程中，将该地址写入网卡的只读存储器(EPROM)。因此，插入这块网卡的计算机的 Ethernet 物理地址是 00-60-08-00-A6-38，不管它连接在哪个具体的局域网中，也不管它移动到什么位置，它的物理地址都是不变的，并且不会与世界上任意计算机的 Ethernet 物理地址相同。需要注意的是：MAC 地址 00-60-08-00-A6-38 经常会写为 00600800A638。

4. 关于全局管理/本地管理(G/L)位与单播/多播(I/G)位的规定

图 4-12 给出了关于全局管理/本地管理(G/L)位与单播/多播(I/G)位的示意图。图中物理地址是 00-60-08-00-A6-38 的二进制写法是按照 IEEE 802.3 规定，将每一个字节的最低位写在左边。

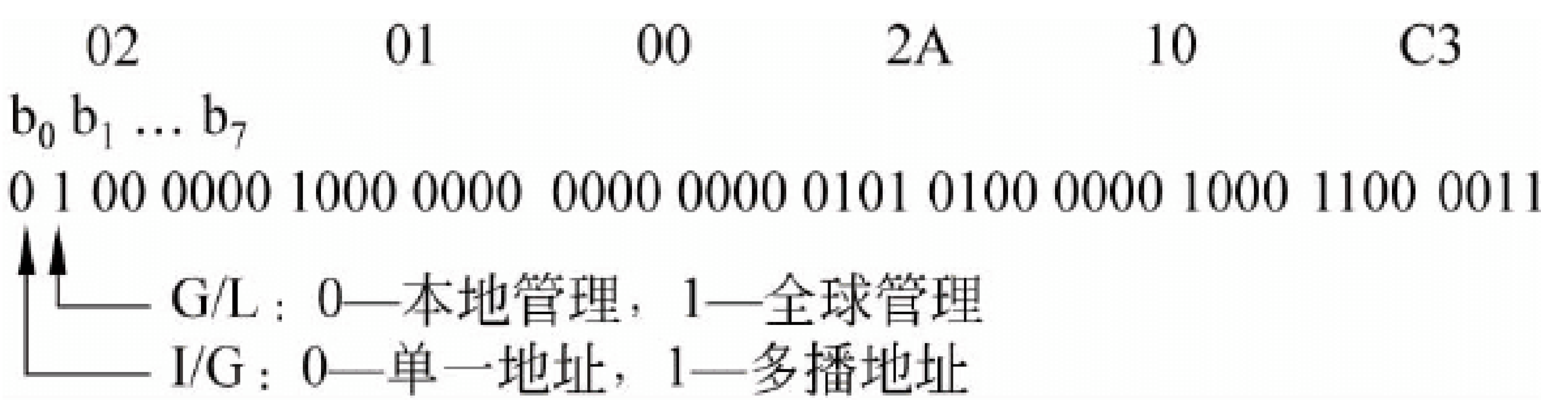


图 4-12 G/L 位与 I/G 位的规定

(1) 在 IEEE RAC 初期制定 Ethernet 物理地址时，考虑到如果有人不愿意向 IEEE RAC 购买 OUI，因此就规定 Ethernet 物理地址的第一个字节的最低的第二位为 G/L (Global/Local)，如图 4-12 所示。G/L＝0 表示本地管理的物理地址，用户可以任意分配，但是不能够保证这个地址是全球唯一的；G/L＝1 表示全局管理的物理地址。所有向 IEEE RAC 购买 OUI 的 Ethernet 物理地址 G/L＝1。事实上，只有个别令牌环网采用 G/L＝0 的本地管理的物理地址。所有 Ethernet 网卡的 MAC 地址都是 G/L＝1 的全局管理的地址，保证在全世界都是唯一的。

(2) 在 IEEE RAC 初期制定 Ethernet 物理地址时，考虑到单播地址、多播地址与广播地址的区别，规定 Ethernet 物理地址的第一个字节的最低的第一位为 I/G(Individual/Group)。I/G＝0 表示单播地址；I/G＝1 表示多播地址。在实际应用中，构造 Ethernet 帧的软件在目的地址字段可能使用的是单播地址、多播地址，或者是广播地址。接收端可以根据 I/G 值来判断，该帧是单播或是多播。

48 位全 1(FF-FF-FF-FF-FF-FF)的 MAC 地址是广播地址。有一些网络系统软件，如 NetWare 服务器软件每 60s 发出一个服务公告帧，用来通知网上所有的结点该服务器的存在。服务公告帧的目的地址为广播地址 FF-FF-FF-FF-FF-FF，Ethernet 上所有的结点都需要接收该帧。

问题 4-18：如何设计 Ethernet 网卡？

网络接口卡(Network Interface Card，NIC)又称做网卡，它是将计算机或其他设备连接到局域网的硬件。对于联网计算机来说，网卡被插入主机的 I/O 通道，并作为主机的一个外部设备来工作。在这点上，网卡与其他 I/O 设备卡，如显示卡、异步通信接口适配器卡没有本质的区别。网卡在主机 CPU 的控制下进行数据的发送和接收。

1. Ethernet 网卡电路结构

Ethernet 网卡的主机接口端插入主机的 I/O 总线通道。对于主机来说，网卡是它的一

个外设。从这个角度来看,Ethernet 网卡与主机的其他外设的地位相同。主机与网卡通过控制总线来传输控制命令与响应,通过数据总线来发送与接收数据。主机通过地址总线和控制总线,依据地址与中断号 INT,去识别网卡和网卡中的寄存器,写入或读出命令或响应。Ethernet 网卡内部由三个部分组成:发送电路、接收电路与介质访问控制电路。图 4-13 给出了典型的 Ethernet 网卡电路结构。

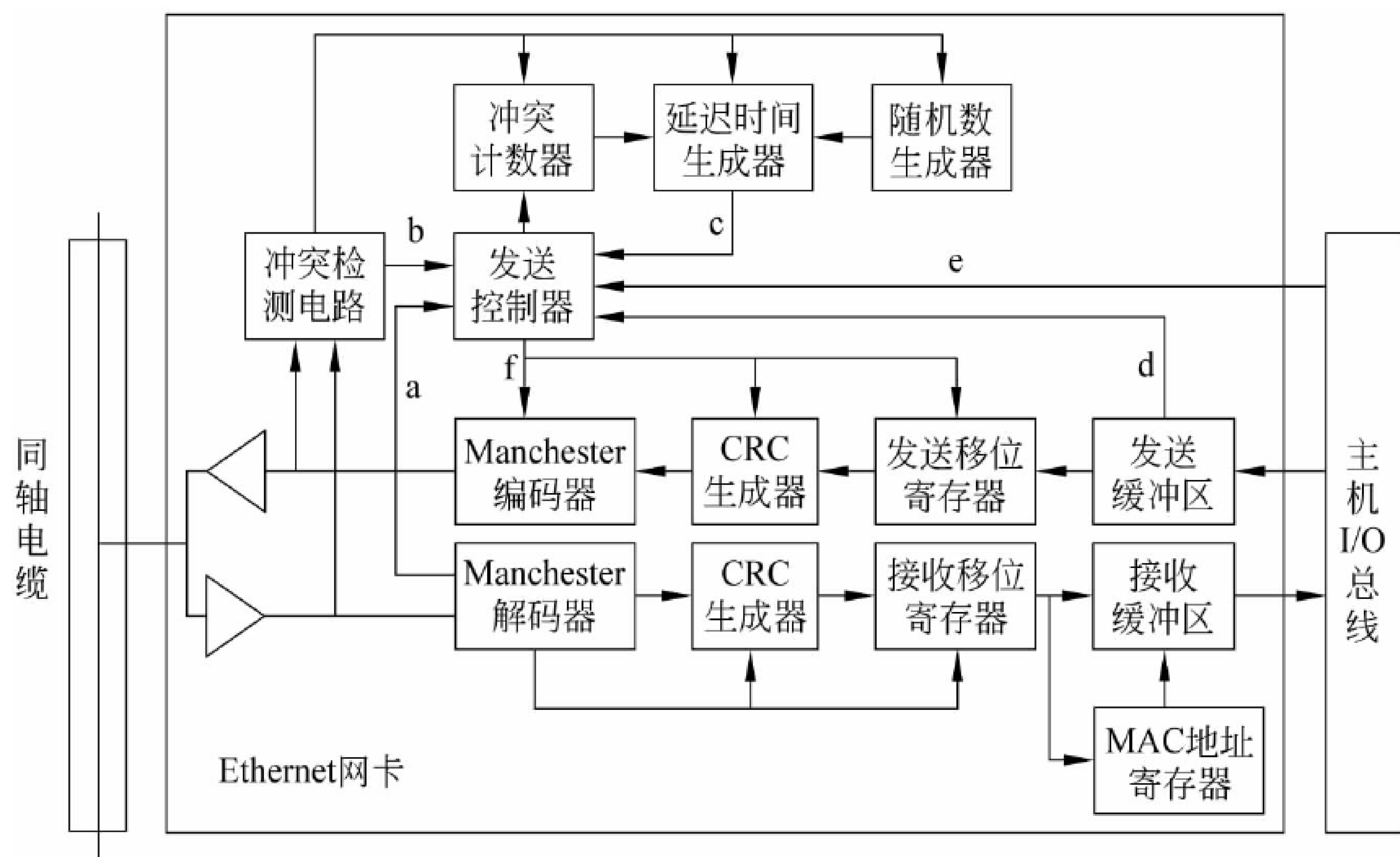


图 4-13 Ethernet 网卡电路结构

2. 执行发送流程

当主机中有数据需要发送时,通过数据总线向输出缓冲区写入数据的同时,向网卡的发送控制器发出发送命令(图 4-13 中的信号 e)。数据通过发送移位寄存器变成串行比特流,通过 CRC 校验生成器与 Machester 编码器后准备发送。根据 CSMA/CD 工作原理,帧是否能够发送取决于传输介质是否忙,需要由发送控制器根据以下几路信号来判断。首先是图 4-13 中的信号 a,它是 Machester 解码器的时钟输出。如果传输介质上有数据信号传输,对应就会产生接收时钟信号。如果接收时钟为 0,说明传输介质上没有数据信号发送,则可以判断“总线空闲”。

如果在发送过程中出现了冲突,冲突检测电路就会向发送控制器发出“冲突发生信号”(图 4-13 中的信号 b)。同时,冲突检测电路向冲突计数器发出控制信号,使冲突计数器、延迟时间发生器与随机数发生器执行“后退延迟算法”。当延迟时间到达时,延迟时间发生器将向发送控制器发出延迟时间到的“重发指示信号”(图 4-13 中的信号 c)。发送控制器在综合考虑“总线空闲信号 a”“冲突发生信号 b”“重发指示信号 c”,以及主机“发送指示信号 e”与“输出缓冲区准备好信号 d”的状态后,将会发出“可以发送信号 f”,它用于控制发送移位寄存器、CRC 生成器与 Machester 编码器。

3. 执行接收流程

从图 4-13 中可以看出,无论是处于发送状态还是接收状态,网卡的接收电路都在接收总线和传输介质中出现信号。在网卡处于发送状态时,接收电路的 Machester 解码器的时钟输出用于判断“总线忙闲”,以及冲突检测电路比较信号的一路输入。

在网卡处于接收状态时，接收到的信号经过 Machester 解码器、CRC 校验器，送到移位寄存器将串行信号变换成并行数据，然后被写入接收缓冲器。在这个过程中，冲突检测电路需要将帧中的目的地址信息与网卡 EPROM 中的本地 MAC 地址进行比较。如果地址匹配，使用"通道注意 CA"信号通知主机接收数据；否则，立即丢弃接收的帧。

4. 支持 CSMA/CD 的 VLSI

在激烈的市场竞争中，一种局域网技术能否大量推广应用，取决于它的性能价格比。采用 CSMA/CD 的 Ethernet 技术的成功，恰恰得益于优秀的技术与开放的市场策略。

1980 年，Xerox、DEC 与 Intel 等三家公司合作，第一次公布了 Ethernet 的物理层和数据链路层规范。1981 年，Ethernet V2.0 规范公布。IEEE 802.3 标准是在 Ethernet V2.0 规范的基础上制定的。实际的网卡均采用可以实现介质访问控制、CRC 校验、曼彻斯特编码与解码、收发器与冲突检测功能的专用 VLSI 芯片构成。

1982 年，第一片支持 IEEE 802.3 标准的超大规模集成电路 VLSI 芯片——Ethernet 控制器问世。此后，很多芯片制造商(例如 Intel、Motorola、AMD 和富士通公司)都能生产 Ethernet 控制器与系列配套芯片。例如，用 Intel 公司的 Ethernet 协处理器 82586，配合 Ethernet 串行接口 82501 和 Ethernet 收发器 82502 芯片，就可以很方便地构成 Ethernet 网卡。Ethernet 协处理器 82586 可以独立于主机的 CPU，完成所有 CSMA/CD 控制功能，并能为开发人员提供高级命令接口。它可以实现 CSMA/CD 介质访问控制、组帧和拆帧、源地址产生、目的地址检查、CRC 校验码生成与校验等功能。82586 提供了两个独立的 16 位 FIFO 来存储接收与发送数据，使用芯片上的 DMA 控制器可以管理网卡与主机的数据交换。网卡与主机最大数据交换速率可以达到 32Mbps。同时，82586 还提供了片上的故障诊断与网络管理功能。Ethernet 串行接口 82501 完成 Machester 编码/解码功能，并且提供与收发器的接口。Ethernet 收发器 82502 完成与传输介质的接口，实现了发送、接收比特流和冲突检测的功能。图 4-14 给出了用 Intel 公司 Ethernet 协处理器 82586，配合 Ethernet 串行接口 82501 和 Ethernet 收发器 82502 芯片设计的 Ethernet 网卡结构。

为了进一步降低个人计算机的联网成本，Intel 又推出了将局域网协处理器 82586 与串行接口 82501 功能集成在一个芯片的 Ethernet 控制器 82588。它的出现使 Ethernet 网卡结构更简洁，同时造价更低廉。图 4-15 给出了用 Ethernet 控制器 82588 与收发器 82502 芯片设计的网卡结构。

问题 4-19：术语辨析：冲突、冲突域、冲突检测、冲突窗口与冲突避免。

理解这几个术语之间的区别与联系需要注意以下几点。

1. 冲突

由于多个结点使用一条公用的总线作为传输介质来发送数据，因此就有可能出现有两个或两个以上结点同时通过总线发送数据，造成接收结点收到的是多个信号电平的叠加，而造成传输失败，这种现象叫作"冲突"。

2. 冲突域

多个结点使用一条公用的总线作为传输介质来发送数据，那么每个时刻只能有一个结点利用总线发送数据，那么我们就说连接在一个缆段上的所有结点共享一个"冲突域"或属于一个"冲突域"。所有通过集线器 Hub 接入局域网的结点，以及通过多个集线器 Hub 级联结构接入局域网的所有结点属于一个"冲突域"。

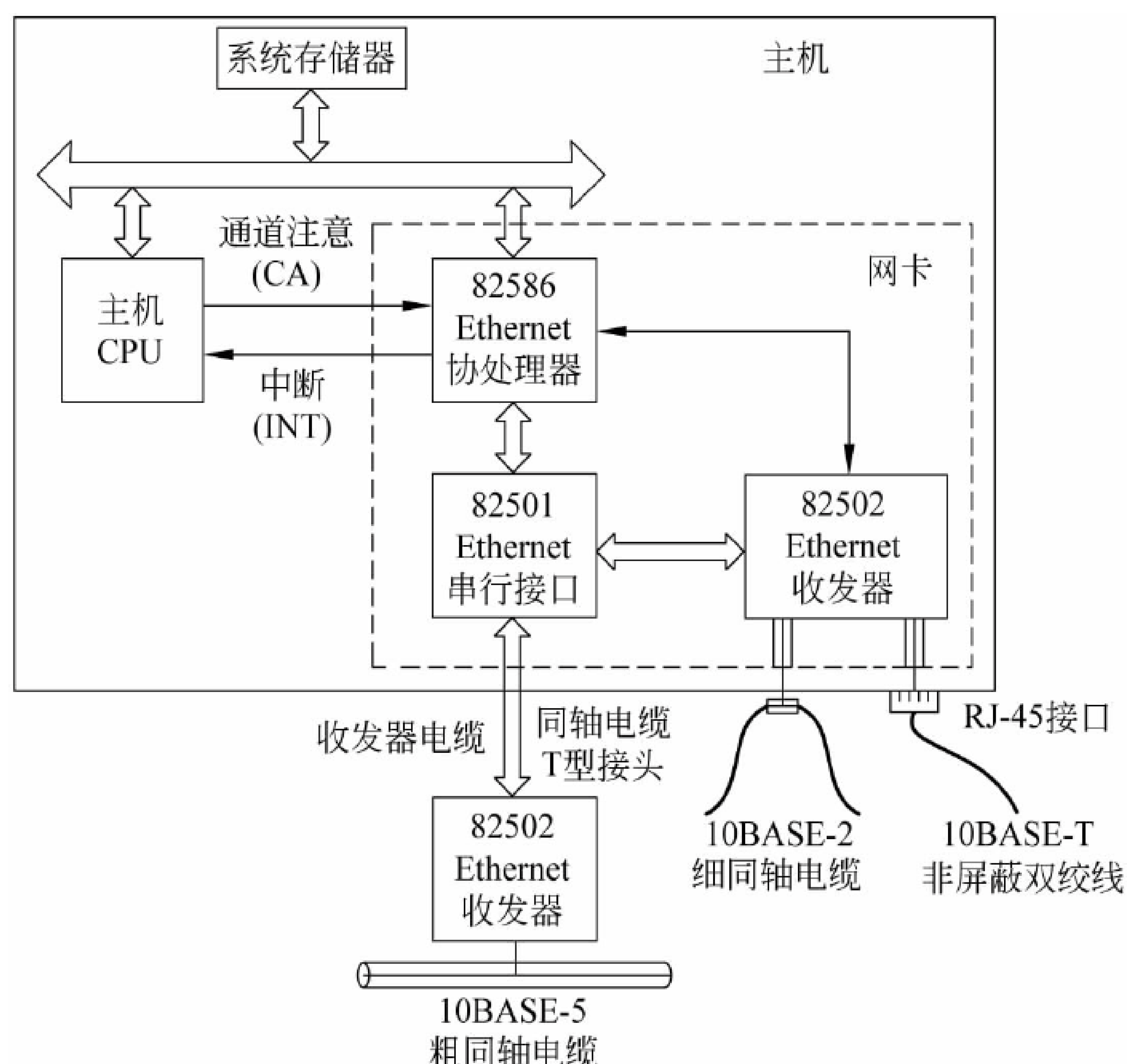


图 4-14　用 Ethernet 协处理器 82586 设计的网卡结构

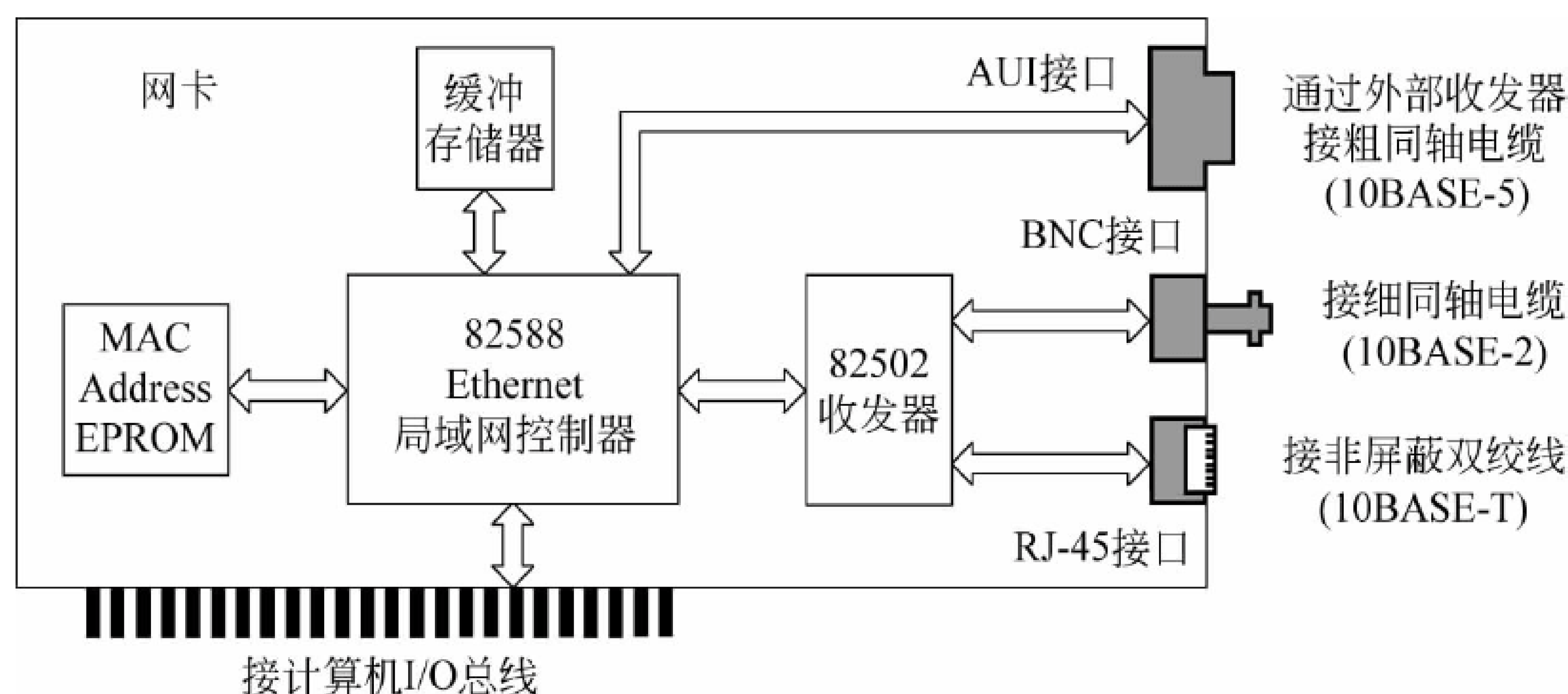


图 4-15　用 Ethernet 控制器 82588 设计的网卡结构

3. 冲突检测

为了确定是否出现冲突而造成传输失败,结点的 MAC 层需要采取一定的方法去发现总线上是不是出现冲突,这种方法叫作“冲突检测”。

4. 冲突窗口

连接在一个缆段上的所有结点都能够检测到冲突发生的最短时间。速率为 10Mbps 的 Ethernet 协议中规定的冲突窗口长度为 51.2μs。Ethernet 的数据传输速率为 10Mbps,冲突窗口的 51.2μs 可以发送 512b(64B)数据。64B 是 Ethernet 的最小帧长度。这意味着当一个结点发送一个最小帧或一个帧的前 64 个字节时没有发现冲突,则表示该结点已经获得总线发送权,并可以继续发送后续的字节。因此,冲突窗口又称为“争用期”。

问题4-20：术语辨析：最小帧长度、最大帧长度与最小帧间隔。

理解这几个术语之间的区别与联系需要注意以下几点。

1. 最小帧长度

IEEE 802.3规定当某个结点完成一帧数据接收后，首先要判断接收的帧长度。如果接收帧长度小于规定的帧最小长度，则表明冲突发生，应该丢弃该帧，结点重新进入等待接收状态。最小帧长度为64B。

2. 最大帧长度

IEEE 802.3规定帧的数据字段最大长度为1500B，那么帧的最大长度为1518B。

3. 最小帧间隔

从接收流程的讨论中可以看出，网卡在接收一帧时需要做一系列的检测和处理。为了保证网卡能正确和连续地处理接收帧，IEEE 802.3标准规定帧间最小间隔值。Ethernet帧间最小间隔时间为9.6μs。接收结点可以利用这段时间处理已接收的帧，并准备接收下一帧，或从接收状态转入发送状态。图4-16给出了最小帧间隔时间示意图。

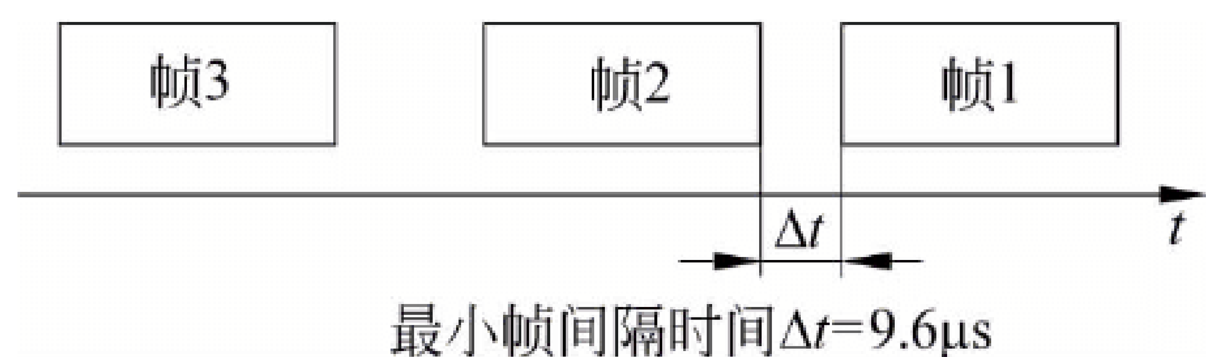


图4-16　最小帧间隔时间示意图

问题4-21：如何理解交换式Ethernet与共享式Ethernet的异同点？

理解这个问题需要注意以下几点。

1. 交互式Ethernet与共享式Ethernet的不同之处

(1) 共享式Ethernet的特点是：MAC控制机制简单，容易实现。存在的问题是连接在一个缆段上的所有结点共享一个"冲突域"，即多个结点共享10Mbps带宽。由于冲突的存在，真正能够利用的总线传输带宽低于10Mbps，并且随着结点数的增加，冲突发生的概率增加，总线的带宽利用率会进一步降低。这样就会造成Ethernet性能的下降。共享式Ethernet采用的是半双工工作方式。

(2) 交互式Ethernet正是针对这个问题，采取以交换机代替集线器；以交换机的并发连接取代共享总线的方式；采用全双工方式代替半双工方式；以独占方式取代共享方式；由于不存在冲突，不采用CSMA/CD控制方法，以提高Ethernet局域网的性能。

2. 交互式Ethernet与共享式Ethernet的相同之处

为了保持与大量部署的传统共享式Ethernet的兼容性，交互式Ethernet保留了传统Ethernet的帧结构、最小与最大帧长度等一些根本的特征。

问题4-22：什么是交换机的线速、线速转发、背板带宽、转发速率与交换带宽？

理解这个问题需要注意以下几点。

1. 线速

线速是指：网络设备（交换机或路由器）的端口转发帧或分组的最大传输速率。例如，Ethernet交换机的每个端口都通过速率为100Mbps的Fast Ethernet网卡与对应主机

100Mbps的Fast Ethernet网卡,通过双绞线连接。我们就可以说交换机的端口线速为100Mbps。这里所说的“线速”是网卡能够发送数据帧的最大发送速率。

2. 线速转发

交换机端口能不能按照“线速”去发送数据帧,不完全取决于端口所能够提供的发送速率的大小,还要看交换机能不能够快速地、及时地处理接收的数据帧。如果交换机不能够及时处理数据帧,势必造成进入交换机的数据帧,因为得不到及时处理而不能够转发,而存放在接收缓冲区中。一旦接收缓冲区全部占用,会因为溢出而造成接收帧的丢弃。因此,“线速”只能说明网卡能够发送数据帧的最大发送速率,对网络设备而言,“线速转发”才是重要的性能指标。“线速转发”是指:接收端无延迟地连续接收帧,交换机能够及时处理,并能连续转发帧。能够以“线速转发”帧表明交换机处于无阻塞,没有出现拥塞的状态。如图4-17所示给出“线速转发”的过程示意图。

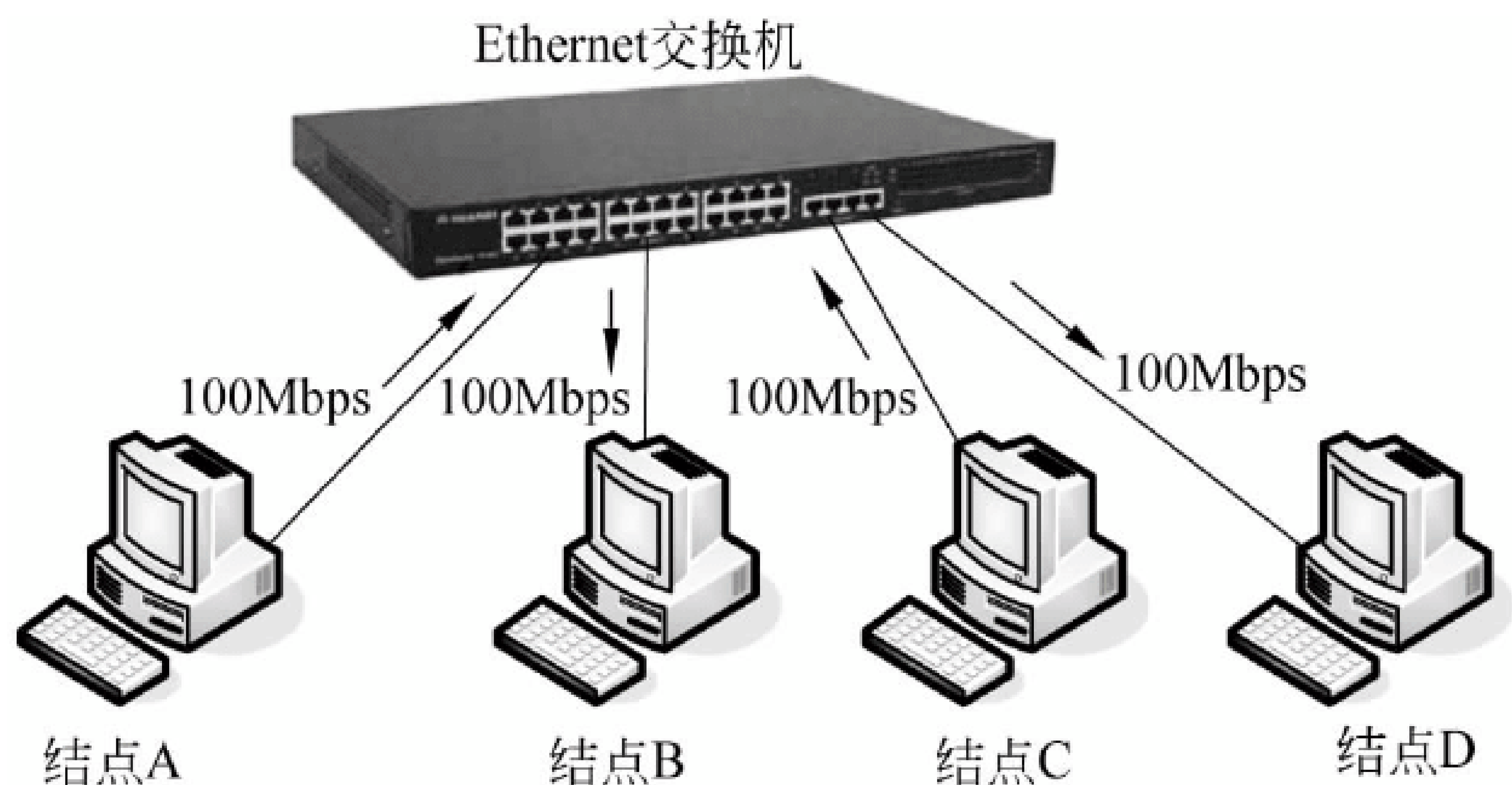


图4-17　线速转发示意图

3. 背板带宽

交换机的背板类似于个人计算机的主板,它担负着帧接收、处理与转发的功能。背板带宽标志了交换机总的数据交换能力,单位为Gbps,叫作交换带宽。交换机是否能够实现“线速转发”,需要看交换机的背板带宽。如果交换机背板交换带宽大于线速交换的要求,交换机才能够实现“线速转发”。

4. 交换带宽

交换带宽的表示有两种方法,一种方法是用交换机的吞吐率来表示,另一种方法是用转发速率来表示。

1) 吞吐率

交换机的交换带宽可以用交换机的吞吐率表示。吞吐率是衡量一个系统或设备单位时间内处理的数据总量。例如,一台交换机有24个100Mbps全双工端口和两个1Gbps全双工端口,如果所有的端口都工作在全双工状态,那么交换机交换带宽为

$$S = 24 \times 2 \times 100\text{Mbps} + 2 \times 2 \times 1\text{Gbps} = 4800\text{Mbps} + 4\text{Gbps}$$
$$= 4.8\text{Gbps} + 4\text{Gbps} = 8.8\text{Gbps}$$

2) 转发速率

转发速率是指:单位时间内交换机能够连续转发帧的数量。按照规定,计算转发速率时是以Ethernet最小长度帧为单位,即帧长度为72B。在讨论接收帧时,不考虑帧前8B的定界符,但是从发送端来看,实际发送的帧长度应该在64B前面加上8B的帧前定界符。同

时，需要考虑帧间间隔（IEEE 802.3 标准规定的帧间间隔为 9.6μs）。

计算的交换机有 24 个 100Mbps 全双工端口和两个 1Gbps 全双工端口，如果所有的端口都工作在全双工状态，那么交换机的转发速率为

（1）Ethernet 的最小长度帧为：72B＝72×8b＝576b。

（2）速率为 100Mbps 时，发送 576b 需要 5.76μs，加上帧间间隔 9.6μs，每帧大约需要用 15.36μs。

每个 100Mbps 端口的转发速率应该为：1/15.36μs≈0.065Mbps。

每个全双工端口的转发速率应该为：0.130Mbps。

其中，1Mbps＝1×10^6（帧/秒）。

24 个全双工端口的转发速率：0.130Mbps×24＝3.120Mbps。

（3）速率为 1Gbps 时，发送 576b 需要 0.576μs，加上帧间间隔 9.6μs，每帧大约需要用 10.136μs。

每个 1Gbps 端口的转发速率应该为：1/10.136μs≈0.099Mbps。

每个全双工端口的转发速率应该为：0.197Mbps。

两个 1Gbps 端口的转发速率应该为：0.197Mbps×2＝0.394Mbps。

（4）交换机的转发速率为：3.125Mbps＋0.394Mbps＝3.52Mbps。

需要指出的是：在估算中没有考虑 GE 对于帧的一些改进方法。

问题 4-23：为什么说虚拟局域网 VLAN 不是一种新型的局域网？

我们在教科书中对虚拟局域网技术的定位是：VLAN 是一种新的局域网服务，而不是一种新型的局域网。理解这个问题需要注意以下几点。

（1）早期的 Token Bus、Token Ring 是可以与 Ethernet 并列的技术，因为它在核心算法上与 CSMA/CD 不同，并且形成了与 Ethernet 的 802.3 标准相并列的标准与产品。

（2）VLAN 不是一种新型的局域网，这一点可以从它的协议上可以看出。1988 年，IEEE 批准了支持 VLAN 的 802.1ac 标准，在 Ethernet 帧的源地址字段与类型字段之间增加了一个 4B 的 VLAN 标记字段。VLAN 标记字段结构如图 4-18 所示。

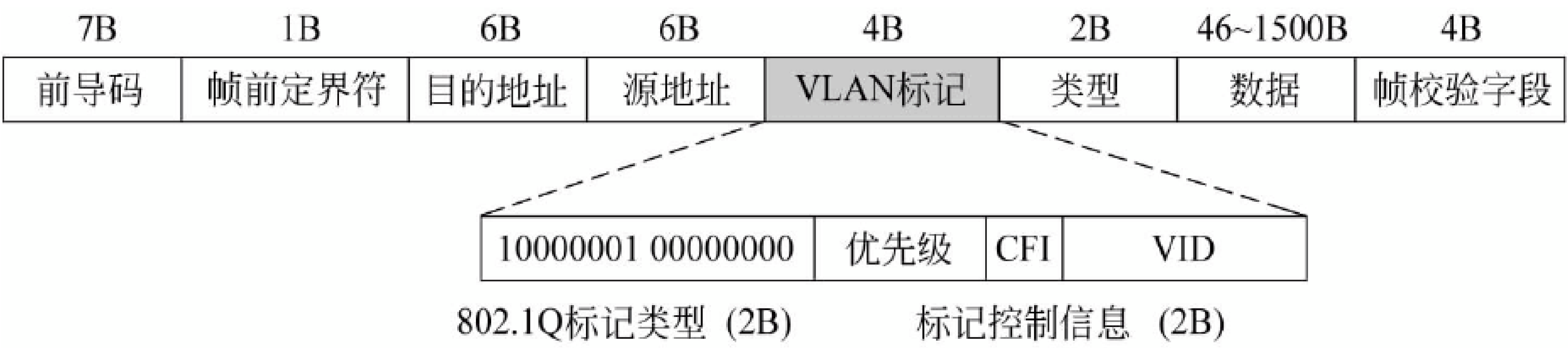


图 4-18　VLAN 标记字段结构

理解 VLAN 标准，需要注意以下几个问题。

（1）VLAN 标记字段组成。

VLAN 标记字段（4B）是由 802.1Q 标记类型字段（2B）与标记控制信息（2B）字段组成。长度 2B 的 802.1Q 标记类型字段值为 0x8100（二进制为 10000001 00000000）。当 MAC 层软件检测源地址字段之后的 2B 值为 0x8100 时，就知道是插入 4B 的 VLAN 标记字段。

(2) 标记控制信息字段。

标记控制信息字段长度是2B,前三位是用户优先级,第4位是规范格式指示符(Canonical Format Indication,CFI)。802标准规范格式是指：地址的十六进制表示中每一个字节的最低位对应规范格式地址中的最低位。CFI=1表示使用地址标识符合802标准的规范格式。

(3) VLAN标识符VID。

VLAN标识符长度为12位,用以标识VLAN的标识VID。VID值表示帧发送的虚拟局域网号。

(4) 由于VLAN的802.1Q标准规定在帧头中增加4B的VLAN标记字段,因此Ethernet帧的最大长度在1518B上增加4B,变为1522B。

IEEE 802.1Q标准是基于Ethernet的802.3协议,目的是在局域网组网中提供更多方便,提高系统安全性。因此,VLAN是一种新的局域网服务,而不是一种新型的局域网。

问题4-24：制定VLAN标准IEEE 802.1Q的难点在哪里?

VLAN研究中的一个重要与困难的问题是：VLAN如何标识? 对于一些新的应用,如IEEE 802.11无线局域网与IEEE 802.16无线城域网,只需要考虑在制定新的协议时,在新的帧头部增加一个字段就可以。例如,IEEE 802.16的连接标识符(Connection ID)与VLAN标识字段的作用是相同的。但是,VLAN的困惑就表现在：它是在已经成熟和广泛应用的Ethernet标准的基础上发展起来,Ethernet帧结构没有为增加VLAN标记字段留任何一点儿余地,而对Ethernet帧结构的改变都会引发一系列的问题。

1995年,IEEE委员会经过多次讨论之后,提出了改变Ethernet帧结构,在Ethernet帧头中增加4B的VLAN标记字段方案。新的格式在1998年公布的IEEE 802.1Q中表现出来。IEEE 802.1Q文档名为：Virtual Bridged Local Area Network。这种改变使得人们不能不思考：已经有那么多的网卡在使用,它们也需要更换吗? 增加4B,那么最大帧长度就不是1518B,而是1522B,软件能够适应吗?

研究结果是：我们不必要求所有连接到VLAN中的主机去接受IEEE 802.1Q协议,而是把执行IEEE 802.1Q协议的工作放在交换机或网桥上。在新型Ethernet交换机上增加IEEE 802.1Q功能是可行的,而不是要去更换Ethernet网卡。同时,在千兆Ethernet(GE)开始考虑与IEEE 802.1Q兼容。GE标准已经将最大帧长度改为1522B。

问题4-25：高速Ethernet采取什么样的发展策略?

(1) 在宽带城域网、三网融合、移动互联网、网络电视、网络视频应用快速发展的基础上,人们不得不寻求有更高带宽的局域网。高速Ethernet的GE、10GbE、40/100GbE技术就是在这种大背景下产生的。

(2) 人们设想了一种用Ethernet组建企业网的全面解决方案：桌面系统采用传输速率为10Mbps的传统Ethernet,部门级网络系统采用传输速率为100Mbps的FE技术,企业级网络系统采用传输速率为1Gbps的GE技术,园区网与城域网接入层主干使用速率为10Gbps的10GbE技术,国家主干网采用40/100GbE技术。由于传统Ethernet与Fast Ethernet、Gigabit Ethernet有很多相同点,并且很多企业已大量使用10Mbps的Ethernet,因此当局域网系统从传统Ethernet升级到100Mbps或1Gbps时,网络技术人员不需要重

新进行培训。相比之下，如果将现有的 Ethernet 互联到作为主干网的 622Mbps 的 ATM 局域网中，一方面由于 Ethernet 与 ATM 工作机制存在较大差异，由于工作机制与协议的不同会出现异型网互联的复杂问题，也就是一种局域网发送的帧格式必须经过转换才能被另一种局域网接受，这种 ATM over Ethernet 的协议转换必然造成系统性能下降。另一方面，熟悉 Ethernet 技术的人员可能并不熟悉 ATM 技术，则对网络技术人员需要重新进行培训。采用高速 Ethernet 组网策略可以平滑地将 10Mbps 的传统 Ethernet、100Mbps 的 FE，以最小代价升级为 1Gbps 的 GE、10Gbps 的 10GbE，甚至是 100GbE 的大型、宽带局域网。随着 GE、10GbE、40/100GbE 技术的成熟，它已经成为大、中型网络系统的首选方案。

（3）GE、10GbE、40/100GbE 的设计思想主要表现在：希望在传输速率进一步提升的同时，仍保持与现有 Ethernet 标准的兼容。高速 Ethernet 在具体设计中采取以下策略。

① 在 MAC 层不使用 CSMA/CD 方法与半双工方式，采用交互式与全双工工作方式。

② 保持相同的 Ethernet 帧结构与帧的最小、最大长度，只是在物理层做一些必要的调整，定义了新的物理层标准。

③ 设计介质专用接口，向 MAC 层屏蔽物理层采用技术的变化。

④ 传输介质逐步向以光纤为主的方向发展；联网设备采用交换机。

（4）GE、10GbE、40/100GbE 采取向传统局域网覆盖范围的基础上向两端扩展，近距离向几十米之内的机房网（存储区域网 SAN、高性能计算机系统、云计算系统）应用方向发展；远距离向宽带城域网、广域网主干网应用方向发展。

（5）GE、10GbE、40/100GbE 考虑到在宽带城域网和广域网中的应用，在标准制定时充分注意 Ethernet 帧信号远距离传输的要求，因此在物理层实现方法、帧格式、MAC 工作速率及适配策略等方面都做出必要的调整。

（6）目前 10GbE 已经开始进入实用阶段。由于 10GbE 技术的出现，Ethernet 工作范围已从校园网、企业网主流选型的局域网，扩大到城域网和广域网。同样规模的 10GbE 造价只有 SONET 的 1/5，ATM 的 1/10。从 10Mbps 的 Ethernet 到 10Gbps 的 10GbE 都使用相同的 Ethernet 帧格式，可以保护已有的应用软件开发投资，并减小网络培训工作量。

问题 4-26：为什么高速 Ethernet 采用 4B/5B、8B/10B 与 64B/66B 编码方法？

高速 Ethernet 在信号编码方法上都做出了改变，例如，100BASE-T 采用的 4B/5B、1000BASE-LX 采用的 8B/10B、10000BASE-ER 采用的 64B/66B 等。做出这样改变的主要原因是考虑到物理层实现技术存在的问题，在保证收发双方比特同步的前提下，降低发送时钟，提高编码效率，降低实现技术的难度。我们可以用 100BASE-T 为例说明这个问题。

（1）10BASE-T 传输比特流之所以要采用 Machester 编码方法，是因为 Machester 编码方法自含同步时钟。但是，Machester 编码方法的传输速率为 10Mbps，而时钟频率需要达到 20MHz。如果 100BASE-TX 仍采用 Machester 编码方法，则时钟频率就需要达到 200MHz。从电子学实现的角度来看，时钟频率提高到 200MHz，造价将会大幅度上升。

（2）100BASE-T 采用的不是传统的 Machester 编码方法，而是 4B/5B 编码方法。4B/5B 编码方法是将 4 位的数据变换成 5 位的码字。表 4-2 给出了十六进制数 0～F 对应的 4B/5B 编码方案。4B/5B 编码方法将待发送的比特以 4 位作为一组，再按照表 4-1 的对应关系转换成 5 比特。

表 4-2　十六进制数 0～F 对应的 4B/5B 编码值

数据	4B 二进制数	5B 编码
0	0000	11110
1	0001	01001
2	0010	10100
3	0011	10101
4	0100	01010
5	0101	01011
6	0110	01110
7	0111	01111
8	1000	10010
9	1001	10011
A	1010	10110
B	1011	10111
C	1100	11010
D	1101	11011
E	1110	11100
F	1111	11101

(3) 4 位二进制代码可以有 16 种组合,而 5 位二进制代码可以有 32 种组合。从表 4-1 中可以看出,4B/5B 编码可以取其中的 16 种组合表示 0～F,其他组合可以用于物理层传输控制的特殊代码。4B/5B 编码方法早期用于 FDDI 环网中。

(4) 从 32 种组合选择出 16 种,可以满足以下两个规则。

① 每个 5 比特组合中不含有多于三个“0”。

② 或者每个 5 比特组合中不含有少于两个“1”。

(5) 在双绞线上传输 4B/5B 编码信号时,采用的方法是“3 级多电平传输”(3-level Multi-Level Transition,MLT-3)方法。MLT-3 用正、负和零等三种电平来传输信号,其编码规律是:在一个传送时钟中,信号电平发生跳变表示逻辑 1,不发生跳变表示逻辑 0。

图 4-19 给出了 MLT-3 信号的编码过程。在双绞线上传输的电信号波形就是 MLT-3 信号波形。MLT-3 信号带有足够的电平跳变,通过这些跳变提取收发双方的同步信号比较容易,同时电平跳变的频率远小于 200MHz 时钟频率。因此,MLT-3 信号的编码、传输与接收、解码,以及比特同步的电子学实现技术上比较容易解决。

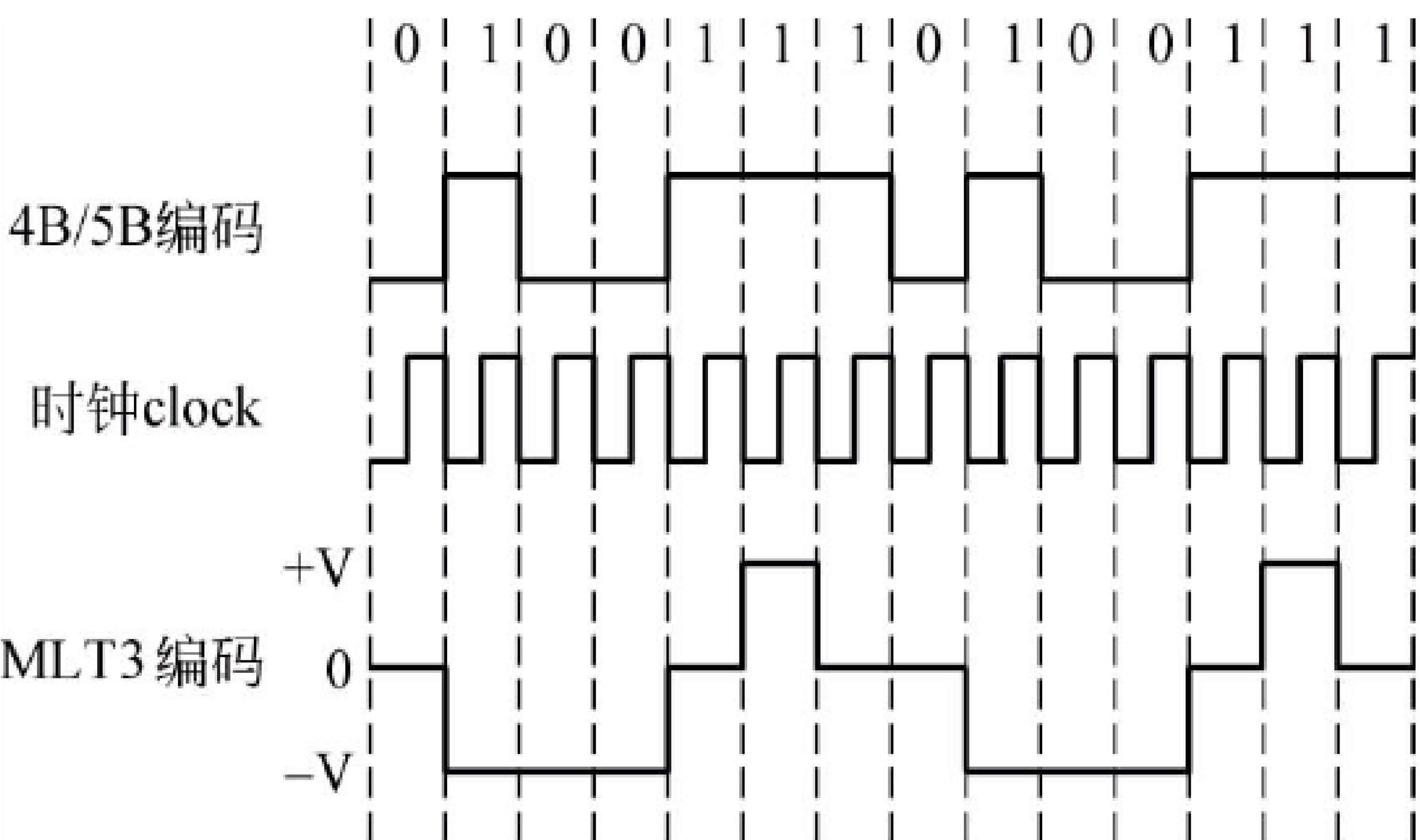


图 4-19　MLT-3 信号的编码过程

(6) 由于采用的方法是“多电平传输”(Multi-Level Transition,MLT)方法,MLT 在一个传送时钟中,信号电平发生跳变表示逻辑 1,不发生跳变表示逻辑 0,因此各种编码中要求尽量多的“1”,尽量少的“0”。

(7) Manchester 编码在每个周期 T 内跳变两次,而只能够传输 1b,因此 Manchester 编码效率为 50%。而 4B/5B 编码效率可以达到 80%。传输速率是 100Mbps,也就是一秒钟发送 1×10^8 个比特。如果采取 Manchester 编码方法,用波特率表示则为 2.0×10^8B。如果采取 4B/5B 编码方法,用 5 位一组的比特组合代替 4 比特来发送,那么用波特率表示则为 $(1\times10^8/4)\times5=1.25\times10^8$B。显然,4B/5B 编码方法的波特率远低于 Manchester 编码方法的波特率。

(8) 更高传输速率的 1000BASE-LX 采用的 8B/10B、10000BASE-ER 采用的 64B/66B 编码的思路与 100BASE-T 的 4B/5B 编码是相同的。

问题 4-27:高速局域网为什么在物理层与 MAC 层之间都增加了介质专用接口 MII?

快速以太网 Fast Ethernet 的 802.3u 标准定义了介质专用接口(MII),802.3z 标准定义了千兆介质专用接口(GMII),10GbE 定义了专用的介质专用接口(10GMII),将 MAC 层与物理层分隔开。图 4-20 给出了快速以太网 802.3u 标准定义的介质专用接口(MII)位置与传统 Ethernet 协议结构的比较。



图 4-20　802.3u 标准的 MII 位置与 Ethernet 协议结构的比较

因为高速 Ethernet 为了保持与传统 Ethernet 的兼容性,它尽管可以不采用 CSMA/CD 随机争用的介质访问控制机制,但是它们必须保持 Ethernet 的帧结构、最大与最小帧长度等基本特征。在物理层要提高传输速率时,必然要在使用的传输介质和信号编码方式方面有所变化。重要的是,高速 Ethernet 在物理层的改变不能够影响 MAC 层,因此需要设计一个介质专用接口(MII)来隔离 MAC 层与物理层。这就是高速 Ethernet 增加介质专用接口(MII)的原因,同时也体现出高速 Ethernet 保持兼容性的设计思想。

问题 4-28:为什么 Fast Ethernet 需要设计速率自动协商机制?

(1) 由于在设计应用中,速率达到 1Gbps 或 10Gbps 的高速 Ethernet 多用于核心层主干网或汇聚层的网络中,只有速率为 100Mbps 的 Fast Ethernet 可能用于接入层,与 10Mbps 的 Ethernet 用户计算机在一个局域网中共存,因此 Fast Ethernet 从与 10Mbps 的 Ethernet 兼容的角度去考虑设计速率自动协商机制。

(2) 速率自动协商应该具有以下功能。

① 自动确定非屏蔽双绞线的远端连接设备使用的是半双工(CSMA/CD)的 10Mbps 工作模式,还是全双工的 100Mbps 工作模式。

② 向其他结点发布远端连接设备的工作模式。

③ 与远端连接设备交换工作模式相关参数,协调和确定双方的工作模式。

④ 自动协商功能自动选择共有的最高性能的工作模式。

(3) 自动协商功能是链路两端设备通过交换100BASE-T定义的“基本链路代码字”来实现的。基本链路代码字长度为16b。图4-21给出了基本链路代码字的结构。基本链路代码字中的主要位的意义是：S0～S4＝00001表示使用的是IEEE 802.3协议；A0位表示10BASE-T半双工，A1位表示10BASE-T全双工，A2位表示100BASE-T半双工，A3位表示100BASE-TX全双工，A4位表示10BASE-T4半双工，A5位表示支持帧流量控制；RF位表示远端故障，ACK位表示确认。

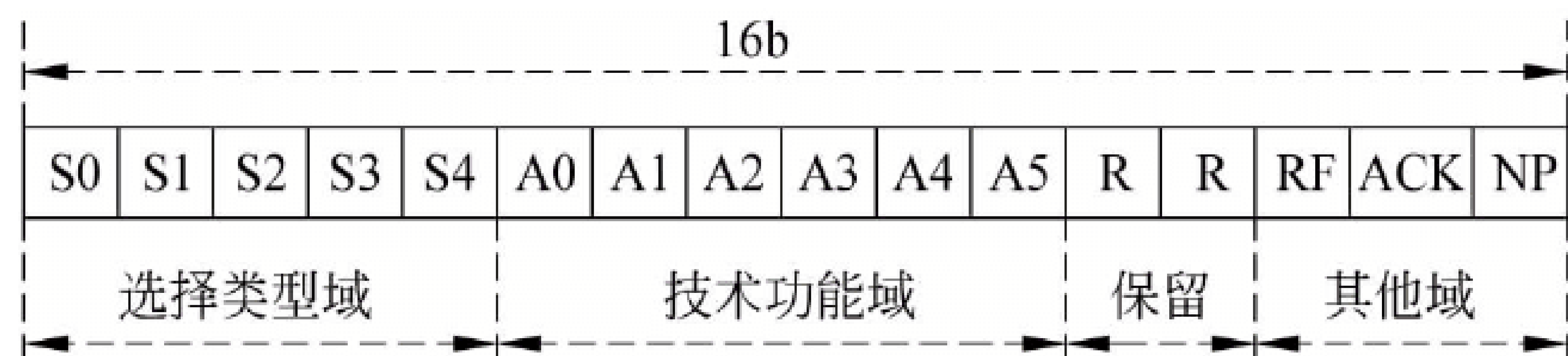


图4-21　基本链路代码字的结构

问题4-29：为什么10GbE帧封装在OC-192帧中传输的数据传输速率不是10Gbps?

回答这个问题需要注意以下几点。

1. 10GbE通过SONET的帧发送与接收过程

10Gbps的Ethernet(10GbE)通过SONET传输网传输时，需要封装在SONET标准的帧封装过程，仅从物理层比特流传输的角度封装成OC-192帧，而对10GbE帧结构不做任何的修改。当发送端的MAC层将帧传送到物理层封装到OC-192帧时，需要增加物理层帧的标识。当接收端的物理层接收到一个OC-192帧后，需要通过拆分OC-192帧还原出原MAC帧中的比特序列，然后将还原的MAC帧比特序列提交给MAC层处理。这个封装与拆分OC-192帧的过程是在物理层进行的，对源结点和目的结点的MAC层是透明的，并非真正修改了MAC帧结构。图4-22给出了10Gbps Ethernet的帧发送过程。

2. OC-192帧的净荷速率

10GbE的广域网物理层应该符合SONET的OC-192/STM-64标准。OC-192的传输速率为9.58464Gbps，而不是精确的10Gbps。在这种情况下，10Gbps的Ethernet帧将被插入到OC-192/STM-64帧的有效载荷中，以便通过光纤系统传输。因此，10GbE的广域网MAC层需要通过10Gbps介质独立子层MII接口来实现9.58464Gbps的速率匹配。如果10GbE的广域网物理层直接采用光纤波分复用DWDM技术，速率为10Gbps。

问题4-30：10GbE的物理层协议有多少种类型?

图4-23给出了10GbE的物理层协议的构成示意图。10GbE的应用领域已经从局域网，逐渐扩展到城域网与广域网的核心交换网中。从图中可以看出以下几个特点。

(1) 10GbE的物理层协议分为两类：局域网物理层标准与广域网物理层标准。

(2) 局域网物理层(LAN PHY)标准：传输介质分为光纤与双绞线两类。

① 基于光纤的物理层协议主要包括以下几个。

- 10GBASE-SR：多模光纤，最大长度为300m。
- 10GBASE-LRM：多模光纤，最大长度为220m。
- 10GBASE-LX4：单模光纤，最大长度为10km。

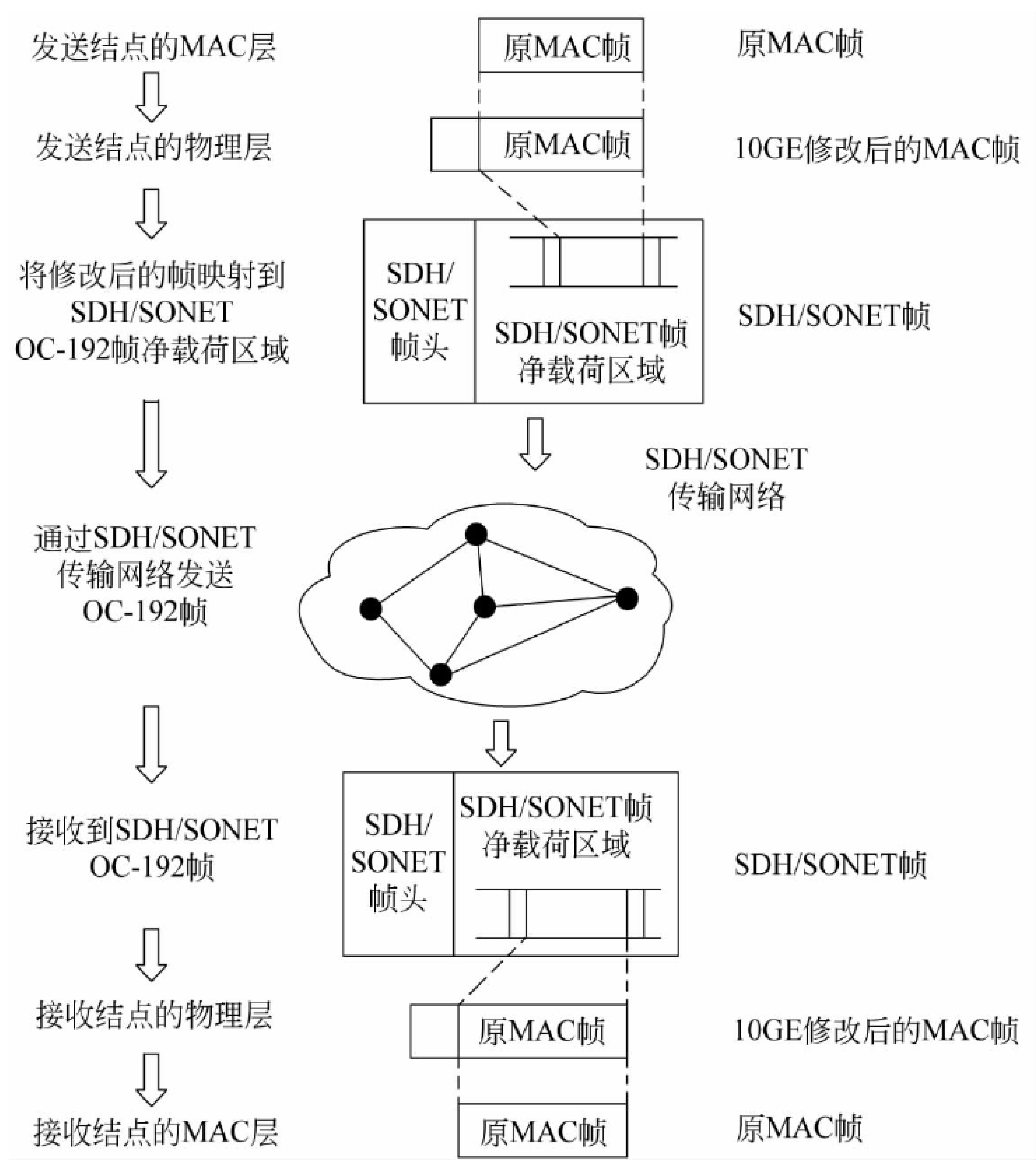


图 4-22　10GbE 帧通过 SONET 传输网的发送与接收过程

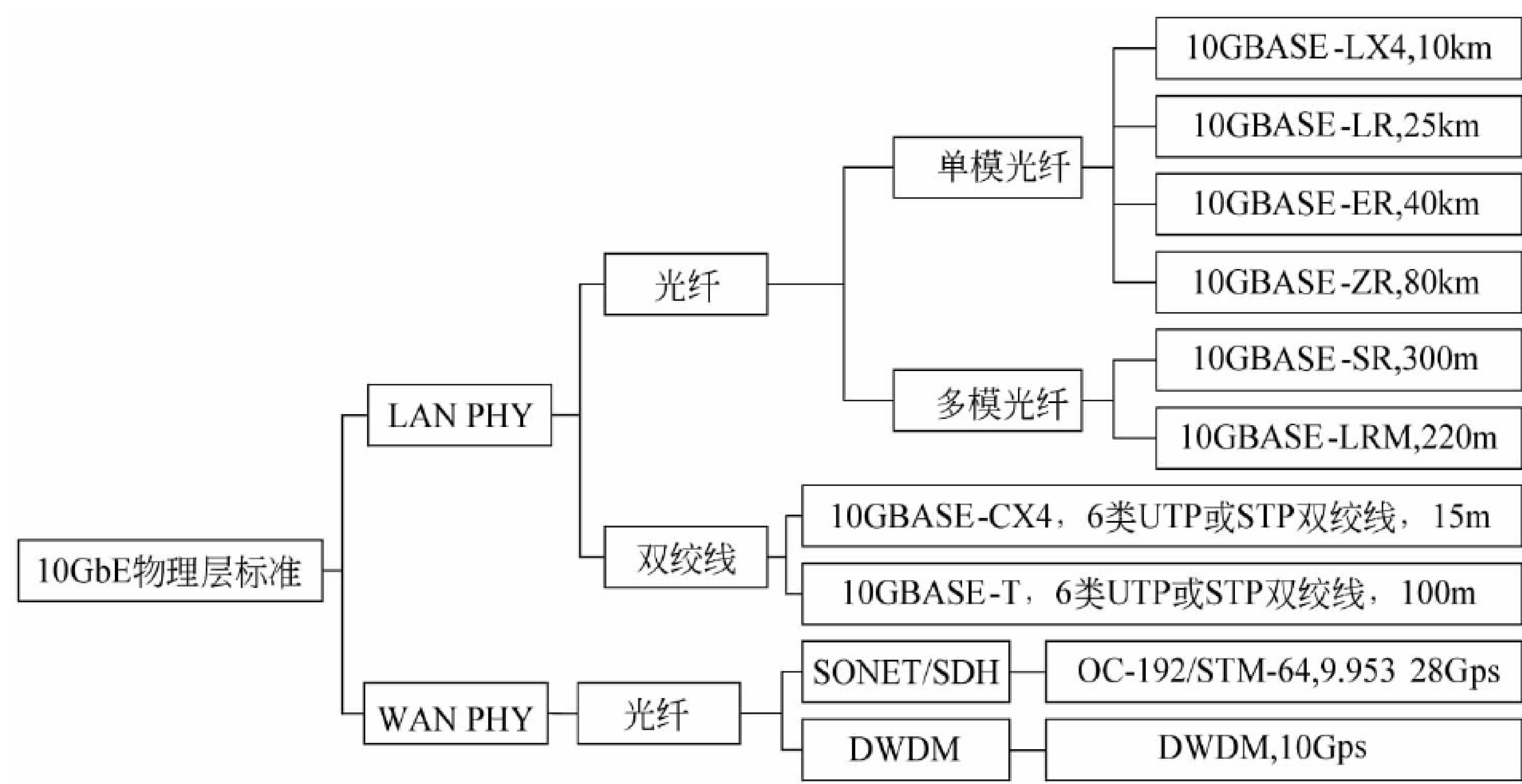


图 4-23　10GbE 的物理层协议的构成示意图

- 10GBASE-LR：单模光纤，最大长度为 25km。
- 10GBASE-ER：单模光纤，最大长度为 40km。
- 10GBASE-ZR：单模光纤，最大长度为 80km。

② 基于双绞线的物理层协议。

基于双绞线的物理层协议主要包括以下几个。

- 10GBASE-CX4：6类UTP或STP双绞线，双绞线最大长度为15m。
- 10GBASE-T：6类UTP或STP双绞线，双绞线最大长度为100m。

(3) 广域网物理层(WAN PHY)标准。

实现WAN PHY标准的技术路线主要有两种：使用SONET/SDH光纤通道技术，以及直接采用光纤密集波分复用DWDM技术。

① 10GbE如果使用光纤通道技术，10GbE广域网物理层应符合光纤通道技术速率体系SONET/SDH的OC-192/STM-64的标准。OC-192/STM-64的标准速率是9.953 28Gbps，而不是精确的10Gbps。

② 直接采用光纤密集波分复用DWDM技术，10GbE速率保持为10Gbps。

问题4-31：如何理解光以太网、城域以太网的特点及它们之间的关系？

回答这个问题需要注意以下几点。

1. 总体认识

(1) 光以太网(Optical Ethernet)与城域以太网(Metro Ethernet)的概念都是在2000年前后提出的，并得到学术界与产业界的认同和支持。

(2) 光以太网与城域以太网标志着Ethernet应用已经从传统的局域网向宽带城域网与广域网领域发展。

(3) 光以太网与城域以太网的概念是密不可分的。光以太网的概念偏重于技术，而城域以太网的概念更偏重于应用。

2. 光以太网的基本概念

1) Ethernet技术的发展

传统的10Mbps的Ethernet的基本特征是：采用双绞线与集线器组网，CSMA/CD介质访问控制方法的半双工的共享传输介质方式，后期也增加使用光纤的10BASE-F标准。100Mbps的Fast Ethernet的基本特征是：采用交换方式与共享方式、全双工与半双工共存的思路，也制定了光纤的100BASE-F标准。1Gbps的Gigabit Ethernet同样保留了交换方式与共享方式共存的思路，但是基于光纤的物理层标准比重增大，并且开始用于宽带城域网的建设。在10Gbps、40Gbps、100Gbps的Ethernet中只采用全双工模式，物理传输介质以光纤为主。由于10Gbps、40Gbps、100Gbps的Ethernet仍然保留着传统Ethernet的帧结构等基本特征，可以保持与大量采用Ethernet技术的网络用户的兼容性。同时，由于不再需要采用CSMA/CD的介质访问控制方法，因此传输介质的长度不需要受冲突窗口的限制。

2) 光以太网的设计思想

研究人员可以充分地将Ethernet技术与SDH、MPLS与DWDM等成熟的光通信技术交叉融合、优势互补，以提升Ethernet技术的服务质量QoS、网络安全性与系统可靠性，使得光以太网成为能够满足电信级服务要求的网络技术。光以太网研究的核心思想是：利用光纤的巨大带宽资源与成熟、广泛应用的Ethernet技术，为网络运营商建造新一代的宽带城域网提供技术支持。

3. 城域以太网的基本概念

1) 城域以太网的发展背景

在传统的城域网领域，电信运营商已经建成了很多网络资源，铺设了大量的裸光纤，建

设了 SDH 环网、帧中继、DDN 专线或 ATM 交换网,网络带宽有 2Mbps、34Mbps、155Mbps、622Mbps、2.5Gbps 或 10Gbps。而要把这些线路资源连到用户端,线路接口标准与技术差异很大,终端设备成本高昂。随着 Ethernet 技术的成熟与广泛应用,将传统的电信传输网技术与 Ethernet 相结合是一条最佳的路径。

如果说宽带城域网选择网络方案的三大驱动因素是成本、可扩展性和易用性,那么选择 Ethernet 技术作为下一代构建宽带城域网的主要技术是非常恰当的。由于 Ethernet 技术成熟、造价低廉,目前世界上已经拥有上亿的用户。Ethernet 具有良好的扩展性,能够容易地实现从 10Mbps 到 100Gbps 的平滑升级,并且能够覆盖从几十米到 100km 的范围。

2) 城域以太网的要求

从构造电信级的宽带城域网的角度来看,传统 10Mbps 的 Ethernet 技术还存在很多的不足。例如,Ethernet 不能提供端-端的包延时和包丢失率控制,不支持优先级服务,不能保证 QoS;不能分离网管信息和用户信息;不具备对用户的认证能力,这就对按时间和按流量计费造成困难。Ethernet 存在这些问题是很容易理解的,因为在初期设计 Ethernet 时,人们只是考虑它如何在局域网环境中工作。

可运营光以太网的设备和线路必须符合电信网络 99.999%的高运行可靠性。它需要克服传统 Ethernet 的不足,并具备以下特征。

(1) 能够根据终端用户的实际应用需求分配带宽,保证带宽资源充分、合理地应用。

(2) 具有认证与授权功能,用户访问网络资源必须经过认证和授权,确保用户和网络资源的安全及合法使用。

(3) 提供计费功能,能及时获得用户的上网时间记录和流量记录,支持按上网时间、用户流量,或包月计费方式,支持实时计费。

(4) 支持 VPN 和防火墙,可以有效地保证网络安全。

(5) 支持 MPLS,具有一定的服务质量保证,提供分等级的 QoS 网络服务。

(6) 能够方便、快速、灵活地适应用户和业务的扩展。

4. 光以太网与城域以太网的影响

研究可运营的光以太网已经不是单一的技术研究,而是提出了城域以太网的解决方案。光以太网、城域以太网的发展将从根本上改变网络运营商规划、建设、管理思想。

问题 4-32:如何理解 Ethernet 物理层标准命名方法?

IEEE 802.3 标准定义 Ethernet 介质访问控制子层与物理层的协议标准。Ethernet 介质访问控制子层统一使用 CSMA/CD 方法和相同的帧结构,但是物理层技术可以是不同的。Ethernet 的物理层标准不同,表示它采用的传输介质、传输速率、传输介质覆盖范围与组网方式不同。在讨论 Ethernet 物理层标准的命名方法时,需要注意以下几个问题。

(1) 传统 Ethernet 的物理层标准的命名方法是:IEEE 802.3 X Type-Y Name。其中,X 表示数据传输速率,单位为 Mbps;Y 表示网段的最大长度,单位为 100m;Type 表示传输方式是基带还是频带;Name 表示局域网的名称。

(2) IEEE 802.3 10BASE-5 表示传输速率为 10Mbps、基带传输、使用粗同轴电缆,最大长度为 500m 的 Ethernet 物理层标准。

(3) IEEE 802.3 10BASE-2 表示传输速率为 10Mbps、基带传输、使用细同轴电缆,最大长度为 200m 的 Ethernet 物理层标准。

(4) IEEE 802.3 10BASE-T 表示传输速率为10Mbps、基带传输、使用双绞线的Ethernet物理层标准。

(5) 当Ethernet的速率提高之后,使用的传输介质可能从非屏蔽双绞线、屏蔽双绞线变成多模或单模光纤,新的物理层标准的命名方法仍然保持不变。这点在高速Ethernet的讨论中可以清楚地看到。

问题4-33:什么是中继器?

回答这个问题,需要注意以下几点。

1. 中继器出现的时期

中继器(Repeater)出现在局域网发展的初期,用于连接同轴电缆,所以现在在实验室中已经见不到了,但是在讨论Ethernet的发展时,尤其是涉及10BASE-2与10BASE-5时,还会遇到类似的问题。按照中继器所能连接的传输介质可以分为多种类型,例如,粗同轴电缆-粗同轴电缆、粗同轴电缆-细同轴电缆、粗同轴电缆-光缆等。中继器在早期的局域网组网中应用广泛。

2. 中继器的作用

数字信号在同轴电缆中传输时有衰减,并且信号波形会发生畸变。传输介质的长度是与信号的衰减与传输延迟相关的。因此,在使用同轴电缆的局域网物理层协议中,必须对单根同轴电缆的最大长度,以及接入的结点数量加以限制。例如,在使用粗同轴电缆的Ethernet网中,专门为它制定了物理层10BASE-5协议。物理层10BASE-5协议规定,粗同轴电缆的一个缆段的最大长度为500m,接入的结点数量的理论值为1024。实际上,由于同时规定每个结点之间的距离至少为5m,因此实际接入的结点数量不会超过100个。在实际使用中存在着两种特殊的需求,一是如果用户需要单个缆段的长度超过500m;二是接入的结点数超过100个,则简单地从10BASE-5协议考虑是不允许的。为了增加Ethernet中同轴电缆的长度,人们设计了中继器(Repeater)这种设备。图4-24给出了用中继器连接两个Ethernet缆段的结构。

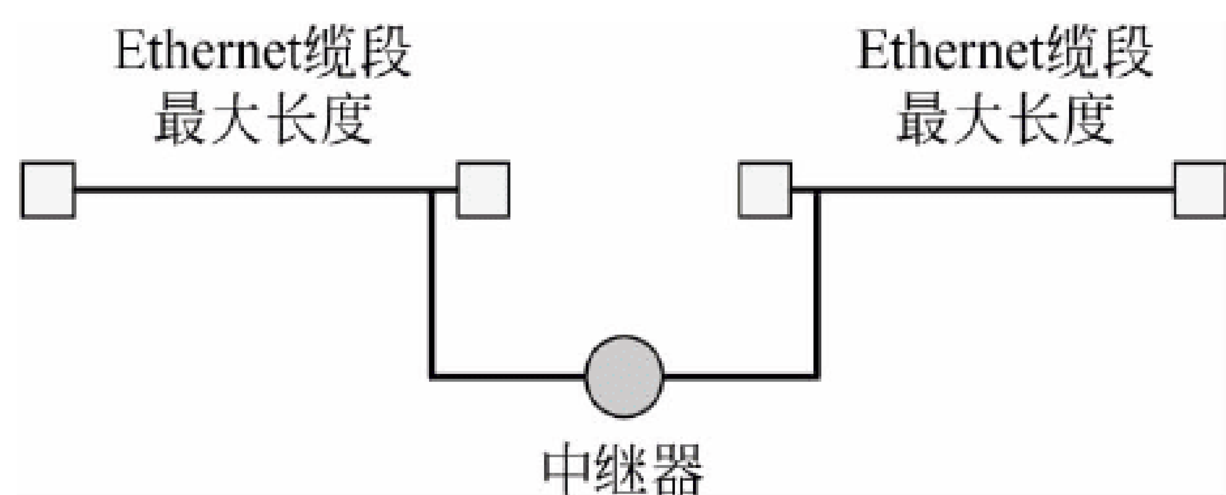


图4-24　用中继器连接两个Ethernet缆段的结构

3. 中继器的基本工作原理

图4-25给出了中继器的工作原理。图中给出的是一个方向的信号传输过程。当位于缆段左侧最远处的主机A发送的信号经过500m同轴电缆的传输后,已经发生了严重的信号衰减和波形畸变。如果发送结点与接收结点之间的距离超过500m,接收结点就不能正确地接收数据信号。因此,设计中继器的目的就是将衰减和变形后的信号,经过接收、放大、整形的工作过程,使得信号的波形与幅度达到协议规定的要求,然后再向它连接的另一个缆段发送出去。

中继器工作在物理层的理由有以下两点。

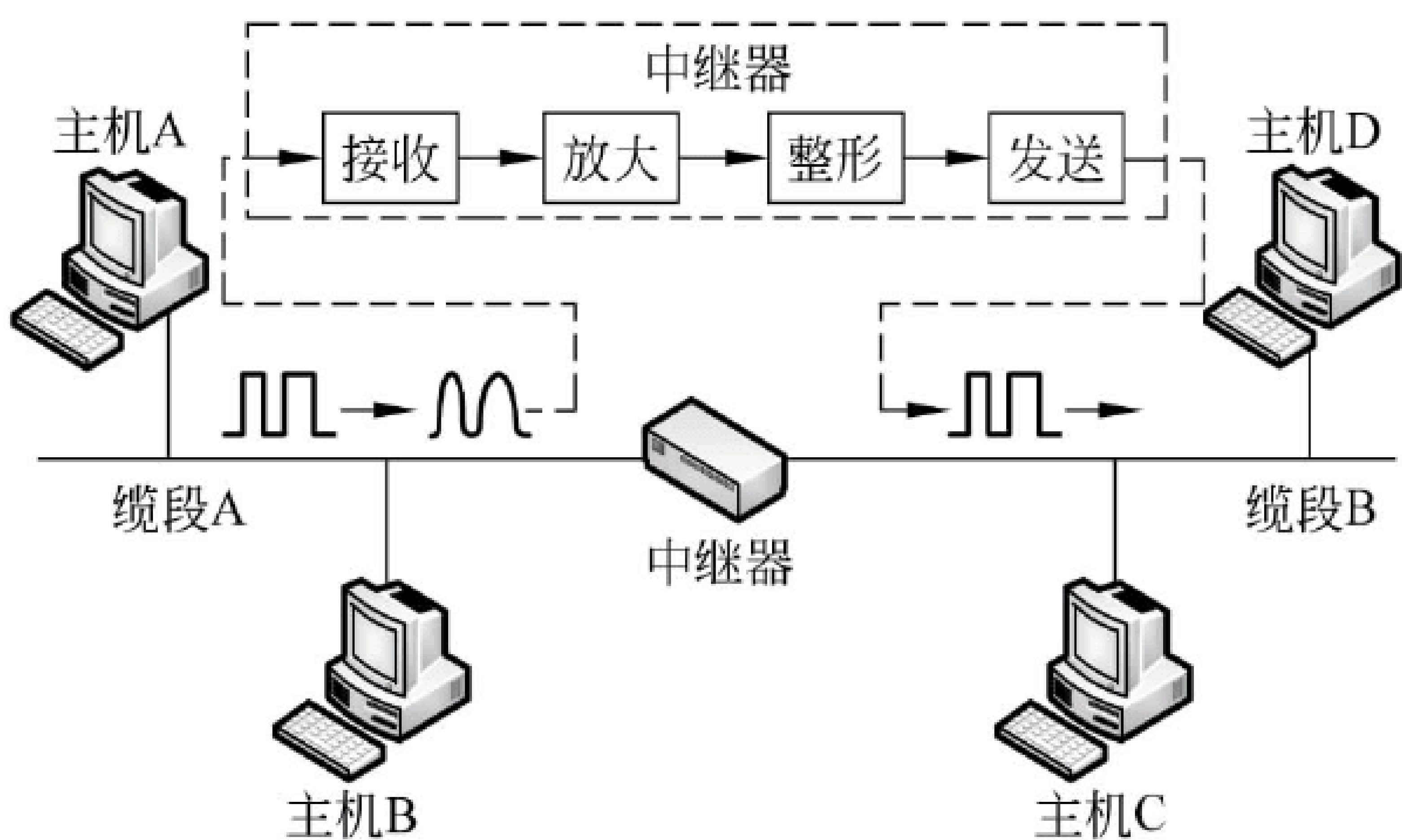


图 4-25　中继器的工作原理

(1) 中继器只能起到对传输介质上信号波形的接收、放大、整形与转发的作用,这是物理层的功能。

(2) 中继器的工作不涉及帧的结构,不对帧的内容做任何处理。中继器只能起到增加传输介质长度的作用。中继器连接的几个缆段仍属于一个局域网。连接在不同缆段上的多个结点,只要有一个发送数据,所有的结点都可以接收到,这些结点共享一个冲突域。因此,中继器不属于网络互联设备。

4. 中继器使用的限制条件

考虑到传输介质的长度与信号的衰减、传输延迟的相关性,因此在一个局域网中使用中继器连接多个缆段的数量是有限制的。例如,在 10BASE-5 协议中,规定最多可以使用 4 个中继器,连接三个缆段,网络中两个结点的最大距离为 2800m。这些规定在 10BASE-5 协议说明中提供了计算依据。

问题 4-34:什么是集线器?

1. 集线器出现的时期

早期的 Ethernet 组网中主要使用粗同轴电缆与细同轴电缆,因此使用中继器比较多。随着 10BASE-T 协议的出现,使用廉价的非屏蔽双绞线 UDP 与 RJ-45 接口就可以实现 10Mbps 的数据传输速率,该技术推动了 Ethernet 的广泛应用。在使用 10BASE-T 协议组网时,集线器的作用就显得十分重要。

2. 集线器的特点

集线器具有以下几个特点。

(1) Ethernet 是典型的总线型结构,设计 CSMA/CD 介质访问控制方法就是在共享总线传输介质的结构下讨论的。当它的物理层采用 10BASE-T 协议时,所有的结点都通过双绞线连接到一个集线器上,它们仍然执行 CSMA/CD 介质访问控制方法,但是从外部结构看其物理结构是星状的。

(2) 所有的结点通过双绞线连接到一个集线器上,它们仍然执行 CSMA/CD 介质访问控制方法,当一个结点发送数据时,所有的结点都能够接收到,因此集线器工作在物理层。图 4-26 给出了冲突域的概念。连接到一个集线器的所有结点共享一个冲突域。

(3) 一个集线器中有多个端口,例如 4 端口、8 端口、16 端口或 24 端口。每个端口通过一个 RJ-45 插头与网卡连接。在 10Mbps 的传输速率下,使用非屏蔽双绞线。标准的非屏

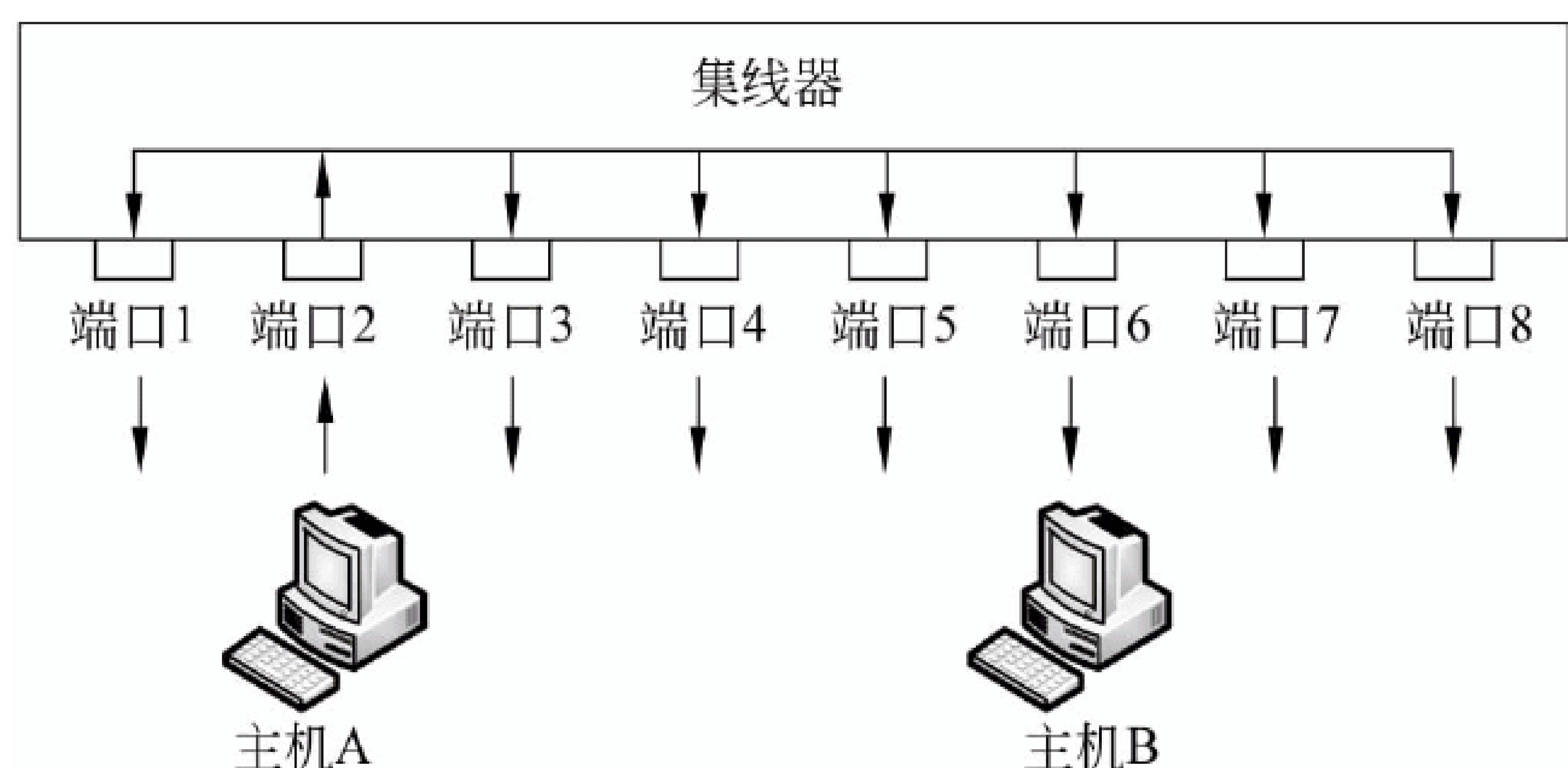


图 4-26　冲突域的概念

蔽双绞线内有4对线,在与网络的连接中实际使用了两对,其中一对用于发送,另一对用于接收。当一个结点需要发送数据时,它将执行CSMA/CD介质访问控制方法,在获得发送数据的权利后,它利用一对发送线将数据通过集线器内部的总线广播出去。

3. 集线器的分类

根据不同的分类方法,集线器可以分为以下不同的类型。

(1) 按集线器支持的传输速率分类,可以分为以下三类:10Mbps集线器、100Mbps集线器、10/100Mbps自适应集线器。

最简单的集线器支持10 BASE-T标准,每个端口的传输速率为10Mbps;目前大多数集线器支持100BASE-T标准。

(2) 按集线器是否能够堆叠,可以分为以下两类:普通集线器、可堆叠式集线器。

普通集线器不具备堆叠功能,当联网结点数超过单一集线器的端口数时,只能采用多集线器的级联方法来扩充;堆叠式集线器由一个基础集线器与多个扩展集线器组成,通过在基础集线器上堆叠多个扩展集线器,可以很方便地扩充联网的结点数。

(3) 按集线器是否支持网管功能,可以分为以下两类:简单集线器、带网管功能的智能集线器。

简单集线器不支持网管功能,无法从远程工作站进行管理;带网管功能的智能集线器支持网管功能,可以通过SNMP来远程监控与管理。

问题4-35:什么是网桥?

理解这个问题需要注意以下几点。

1. 网桥的研究背景

将一个大型局域网划分成多个用网桥或路由器互联的子网,这就导致了局域网互连技术的发展。网桥(Bridge)与路由器(Router)可以隔离子网之间的交通量,使每个子网作为一个独立的小型局域网。通过减少每个子网内部结点数N的方法,可以使每个子网的网络性能得到改善,而每个子网的介质访问控制仍采用CSMA/CD方法。

2. 网桥的应用领域

很多实际应用需要将多个局域网互联,这些应用环境主要有以下几种。

(1) 一个单位的很多部门需要将各自的服务器、工作站与微型计算机互连成网,不同的部门根据各自需要组建独立的局域网,而各个部门局域网之间又需要交换信息、共享资源,这就需要把多个局域网互联起来。

(2) 一个单位有多幢办公楼,每幢办公楼内部建立了局域网,这些办公楼内部的局域网需要互联起来,构成支持整个单位管理信息系统的局域网环境。

(3) 在一个大型的企业或校园内,有数千台计算机需要联网,如果将它们用一个局域网连接,局域网负荷增加会造成网络性能下降。可行的办法是将数千台计算机按地理位置或组织关系划分为多个子网,每个子网是一个局域网,多个局域网互联起来构成一个大型的企业网或校园网。

(4) 如果联网计算机之间的距离超过单个局域网的最大覆盖范围,则可以将它们分成几个局域网来组建,再将这些局域网互联起来构成一个大的网络。

(5) 一个单位的不同部门信息系统的安全性要求与通信量、通信要求是不同的。如果企业某个部门对信息安全、保密要求较高,可以将该部门的计算机连接成独立的局域网,再将这个局域网与其他局域网互联起来,构成一个安全的系统。

3. 网桥的主要特点

网桥在数据链路层完成数据帧接收、转发与地址过滤功能,它用来实现多个局域网之间的数据交换。在使用网桥实现数据链路层的互连时,允许互联网络的数据链路层与物理层协议不同。

网桥是在数据链路层实现网络互联的设备,它具有以下几个基本特征。

(1) 网桥能够互联采用不同数据链路层协议、不同传输介质与不同传输速率的网络。

(2) 网桥以接收、转发与地址过滤的方式实现互联网络之间的通信。

(3) 网桥可以分隔两个网络之间的广播通信量,有利于改善互联网络性能与安全性能。

4. 网桥的基本工作原理

网桥的实现比较简单,在个人计算机中只需选择不同类型的网卡。例如,专门为连接Ethernet与Token Ring设计一种网桥,在主机中插入一块Ethernet网卡与一块Token Ring网卡,这样就可以构成网桥的基本工作环境。实际上这是一个用于异型局域网互联的网桥。因为Ethernet与Token Ring的帧结构、介质访问控制方法与传输速率都不同。在这台计算机上分别装入各种网卡的驱动程序,网卡独立完成各自的帧发送与接收功能。网桥软件完成接收、转发与地址过滤。网卡与网桥软件在一个CPU的控制下完成网桥的基本功能。网桥就可以实现不同的帧结构、介质访问控制方法与不同速率的网络互联。目前,用于异型局域网互联的网桥已不重要,因为局域网主流技术基本都采用Ethernet标准。但是,不同Ethernet技术在物理层可以有不同的速率标准,网桥仍需要解决同种标准中的不同传输速率的局域网在MAC层互联的问题。

5. 网桥的分类

网桥可以有以下几种分类方法。

(1) 根据网桥的帧转发策略来分类,可以分为透明网桥与源路由网桥。

(2) 根据网桥的端口数来分类,可以分为双端口网桥与多端口网桥。

(3) 根据网桥的连接线路来分类,可以分为普通局域网网桥、无线网桥与远程网桥。

问题4-36:什么是透明网桥?

理解这个问题需要注意以下几点。

1. 透明网桥的基本概念

网桥最重要的维护工作是构建和维护路由表。路由表中记录不同结点的物理地址与网桥转发端口关系。如果没有路由表,网桥无法确定帧是否需要转发,以及如何进行转发。IEEE 802.1 与 IEEE 802.5 两个分委员会分别制定了透明网桥与源路由网桥的协议标准。透明网桥标准是 IEEE 802.1d。

2. 透明网桥的主要特点

透明网桥主要有以下几个特点。

(1) 透明网桥由每个网桥自己来进行路由选择,局域网上的各结点不负责路由选择,网桥对于互联局域网的各结点是“透明”的。

(2) 透明网桥一般用于两个 MAC 层协议相同的网段之间的互联,例如,连接两个 Ethernet 网或两个令牌环网。

(3) 透明网桥的最大优点是容易安装,它是一种即插即用设备。

目前,使用最多的网桥是透明网桥。透明网桥的路由表要记录三个信息:站地址、端口与时间。当透明网桥刚连接到局域网时,其路由表显然是空的。当透明网桥接收到一个帧时,它将记录接收帧的源 MAC 地址、帧进入该网桥的端口号与时间,然后将该帧向所有其他端口转发。网桥在转发过程中逐渐建立起路由表。

问题 4-37:什么是源路由网桥?

IEEE 802.5 分委员会制定了源路由网桥标准。源路由网桥由发送帧的源结点负责路由选择。源路由网桥假定每个结点在发送帧时,都已经清楚到目的结点的路由,因此在发送帧时将详细的路由信息放在帧的首部。

问题的关键是:源结点如何知道应该选择的路由。为了发现适合的路由,源结点以广播方式向目的结点发送一个用于探测的发现帧。发现帧在通过网桥互连的局域网中沿着所有可能的路由传输,并在传输过程中记录所经过的路由。当这些发现帧到达目的结点后,将会沿着各自的路由返回源结点。源结点在得到这些路由信息后,从所有可能的路由中选择出一个最佳路由,一般选择经过的中间网桥的跳步数最少的路由。

此后,所有从这个源结点向该目的结点发送的帧首部,都必须携带源结点确定的路由信息。发现帧的另一个作用是帮助源结点确定整个网络可以通过的帧最大长度。

问题 4-38:如何理解生成树协议 STP 的基本内容?

1. 生成树协议 STP 的基本概念

局域网的拓扑经常会发生变化。为了使路由表能反映整个网络的最新拓扑,还需要记录每个帧到达网桥的时间,以便在路由表中保留网络拓扑的最新状态信息。网桥中的端口管理软件周期性地扫描路由表,只要是在一定时间(例如几分钟)以前登记的都要删除,这样就使路由表能反映当前网络拓扑状态。

在很多实际的网络应用中,很难保证通过网桥互连的网络不出现环状结构。环状结构可能使网桥反复转发同一个帧,从而增加了网络中不必要的负荷,进而降低系统性能。为了防止出现这种现象,透明网桥使用了生成树算法。

2. 生成树协议 STP 的基本内容

根据生成树算法制定的协议称为生成树协议(Spanning Tree Protocol,STP)。生成树

协议可以从网络拓扑中清除数据链路层的环路。IEEE 802.1d 规定了 STP,它能够阻断网络中存在的冗余链路,以消除路径中的环路,并可以在活动路径出现故障时,重新激活冗余链路来恢复网络的联通性,保证网络的正常工作。STP 规定了一种特殊的桥协议数据单元(Bridge Protocol Data Unit,BPDU),并使用了以下几个重要的概念:根网桥、最短路径开销、指定网桥、根端口、指定端口与阻塞端口。

根网桥是从网络中选择的一个作为属性拓扑的树根;最短路径开销是一个网桥到根网桥的最短路径;指定网桥负责转发到根网桥的数据;对于每个非根网桥,都需要从它的端口中选出一个到达根网桥路径最短的端口作为根端口,根端口一般处于转发状态;对于每个网段需要选择一个距离根网桥最近的端口作为指定端口,负责将本网段的数据发送到根网桥,这个端口就叫作指定端口,一个网段中只有一个指定端口;生成树协议为每个网段选择一个指定端口,那么其他的端口均处于阻塞状态,因此就叫作阻塞端口。

构造生成树首先要选择一个网桥作为生成树的根。实现方法是选择 ID 最小的网桥作为根网桥。接着,按照根到每个网桥的最短路径来构造生成树。如果某个网桥或局域网失败,则重新计算。该算法的结果是建立起从每个局域网到根网桥的唯一路径。该过程由生成树算法软件自动产生;拓扑结构变化时将更新计算生成树。

3. 桥协议数据单元 BPDU 结构

生成树算法通过网桥之间的一系列协商构造出一个生成树。这些协商的结果是:每个网桥都有一个端口被置于转发状态,而其他端口则被置于阻塞状态。该过程保证网络中任何两个设备之间只有一个通路,可以防止出现任何形式的环路,创建了一个逻辑上无环路的网络拓扑结构。实行生成树算法协商过程的是桥协议数据单元 BPDU。BPDU 的结构如图 4-27 所示。

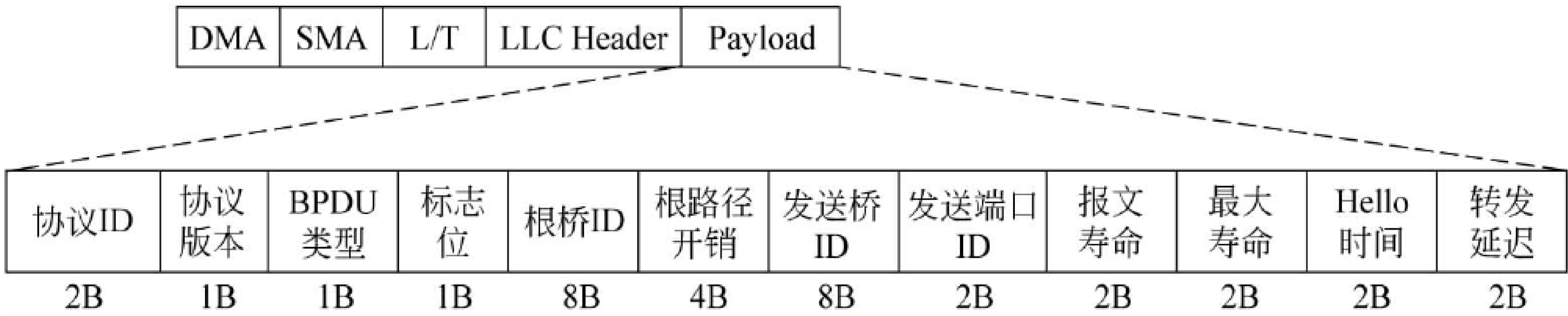


图 4-27 桥协议数据单元 BPDU 结构

在桥协议数据单元 BPDU 中,目的地址 DMA 使用网络上所有网桥都知道的多播 MAC 地址,一旦网桥接收到目的地址为该地址的帧,就能够确定它是桥协议数据单元 BPDU。源地址 SMA 为发送该 BPDU 的网桥 MAC 地址。L/T 表示帧的长度。LLC Header 采用固定的值 0x42。Payload 是 BPDU 的数据,该数据携带着用于生成树计算的所有数据。Payload 中数据最主要的有根网桥 ID、根路径开销、发送桥 ID、发送端口 ID。

根网桥 ID 的 8B 的数值中,前 2B(16b)为根网桥的优先级,后 4B(48b)为根网桥的 MAC 地址。网桥的优先级默认值为 0x8000(十进制数为 32 768),网桥的优先级可以由管理员设定。根路径开销是从发送网桥到根网桥的最短路径开销,即最短路径上所有链路开销的代数和。发送网桥 ID 的 8B 的数值中,前 2B 为网桥的优先级,后 4B 为网桥的 MAC 地址。发送端口 ID 长度为 2B,其值由端口优先级与端口索引号组成。一个 16 端口的网桥,各端口的索引值依次是 1~16;端口的优先级可以由管理员设定。这 4 项参数写成矢量的

形式为：RootBridgeID,RootPathCost,TransmittingBridgeID,TransmittingBridgePathID。

4. 生成树算法的执行过程

生成树算法根据最优配置报文选举的原则运行,其方法如下。

(1) 比较配置报文根网桥ID,根网桥ID小的优先级高。

(2) 如果根网桥ID相等,则比较根路径开销;根路径开销等于接收端口收到的配置报文中的路径开销加上接收端口对应的链路开销,该值小的优先级高。

(3) 如果前面两个参数相等,再比较它们的发送网桥ID,发送网桥ID小的优先级高。

(4) 如果前面三个参数相等,再比较它们的发送端口ID,发送端口ID小的优先级高。

例如,配置报文BPDU1为(12,8,2,12),BPDU2为(12,8,5,8),显然BPDU1与BPDU2根网桥ID与根路径开销是相等的,那么需要比较第三项。BPDU1的根路径开销为2,BPDU2的根路径开销为5,因此BPDU1的优先级高于BPDU2。

每个网桥都有一个最优的配置报文,在网桥加电之后经过短暂的阻塞状态之后,将进入Listening与Learning(默认时间均为15s)这两个过渡状态。在这两个过渡状态中,网桥的每个端口每隔2s都要发送该网桥的最优配置报文,以便选举出根网桥、根端口与指定端口,最后确定该端口是进入转发状态,或者是进入阻塞状态。

问题4-39：什么是广播风暴?

1. 网桥存在的问题

网桥存在着两个主要的问题：帧转发速率低与广播风暴。

评价网桥性能的参数主要有两个：帧过滤速率与帧转发速率。帧过滤速率是指每秒钟能通过多个端口接收并完成帧地址过滤的最大帧数。帧转发速率是指每秒钟能通过多个端口实际转发的最大帧数。由于网桥的帧过滤功能主要由软件完成,因此作为网桥的计算机的CPU速度,对地址过滤和转发速率有着重要的影响。尽管可以从计算机体系结构与软件设计上不断提高网桥的性能,但是如果不从网桥的工作原理的角度与硬件实现帧交换,帧转发速率低的现状只能是有所缓解,而得不到较好的解决。因此,在多个局域网通过网桥互联的结构中,网桥会成为系统性能的瓶颈。

2. 造成广播风暴的原因

从网络体系结构来看,网络系统的最低层是物理层,第二层是数据链路层,第三层是网络层。在讨论网桥的工作原理时,已经知道网桥工作在数据链路层。网桥根据数据帧的源地址与目的地址来决定是否接收和转发该帧。根据网桥的工作原理,网桥对同一子网中传输的数据帧不转发,因此可以达到隔离互联的子网通信量的目的。由于网桥要确定传输到目的结点的帧通过哪个连接端口转发,因此必须在网桥中保存一个“端口-结点地址表”。同时,网桥中保存“端口-结点地址表”的存储器空间有限。因此,随着网络规模的扩大与用户结点数的增加,将会不断出现“端口-结点地址表”中没有的结点地址信息。当带有这种目的地址的数据帧出现时,网桥无从决定应该从哪个端口转发。这时,它的唯一办法就是通过“帧广播”在所有端口广播,只要这个结点在互联的局域网中,广播的数据帧总有可能到达目的结点。这种方法非常简单,但是却带来了很大的问题,那就是一个帧经过一轮又一轮的广播后,变成了2个、4个、8个、16个,这种盲目广播会使帧的数量按指数规律增长,造成网络中无用的通信量剧增,形成“广播风暴”,情况严重时会造成系统无法正常工作。

问题 4-40：局域网网桥与交换机的区别是什么？

交换式局域网的核心设备是局域网交换机，局域网交换机可以在它的多个端口之间建立多个并发连接。为了保护用户已有的投资，局域网交换机一般是针对某类局域网而设计，例如，IEEE 802.3 的 Ethernet 或 IEEE 802.5 的 Token Ring。典型的交换式局域网是交换式以太网(Switching Ethernet)，它的核心部件是以太网交换机(Ethernet Switch)。

实际上，局域网交换机与网桥之间没有严格的界限，可以认为交换机是在网桥的基础上发展起来，并且是功能更加完善的网桥。但是，在结构、交换方式、地址过滤与端口间连接方式等方面，交换机与网桥之间有明显的区别。

1. 结构

网桥通常是在一台典型的个人计算机上配置而成，其内部结构一般只有一个 CPU，通过软件方法完成网桥的接收、存储、地址过滤与转发等功能。因此，网桥的帧过滤速率与转发速率等性能受到结构的限制。为了提高局域网交换机的性能，通常使用针对帧转发设计的专用集成电路芯片 ASIC，或采取多个 CPU 并发工作的计算机结构。

2. 交换方式

网桥一般使用简单的地址过滤与存储转发功能。交换机的交换方式有多种，例如直接交换方式、存储转发交换方式与改进的直接交换方式。

3. 地址过滤与转发策略

网桥的地址过滤与转发策略由一个 CPU 完成。交换机则是在多个端口上并发接收与发送多个帧，从而实现帧的快速转发与地址表的实时更新。

4. 端口间连接方式

网桥的端口是以共享存储空间方式来连接的。交换机则是以硬件方式实现多个端口的并发连接，从而可以大大提高交换机的性能。

5. 端口数量

由于受到自身结构的影响，网桥的端口数量一般比较少，最多不会超过 24 个。交换机的硬件系统是专门设计的，用以实现多个端口的并发连接，因此端口数量比较多，最多可达 128 个。

问题 4-41：中继器、集线器、交换机、网桥、路由器与网关的区别是什么？

回答这个问题需要注意以下几点。

1. 网桥的特点

从网桥工作流程的分析中，可以看出网桥的优点以及存在的问题。

(1) 网桥以接收、存储与转发的方式实现互联局域网之间的通信。网桥将互联的局域网分割成多个冲突域，隔离了局域网之间的流量，改善了互联局域网的性能与安全性。

图 4-28 给出了网桥与中继器作用的比较。由于中继器属于物理层连接局域网的设备，因此它不能够识别 MAC 层地址，它只能够直接将结点 A 发送的数据比特流传播出去，因此由两个中继器连接的三个局域网共享一个冲突域。也就是说，在这三个局域网中，每个时刻只能有一个结点发送数据，其他结点只能够接收数据。网桥工作在 MAC 层，它可以根据帧的源 MAC 地址与目的 MAC 地址来确定接收帧是否应该转发。这样，每一个互联的局域网本身就形成了一个冲突域。因此，网桥可以起到隔离局域网之间的流量，改善局域网的性

能与安全性的作用。

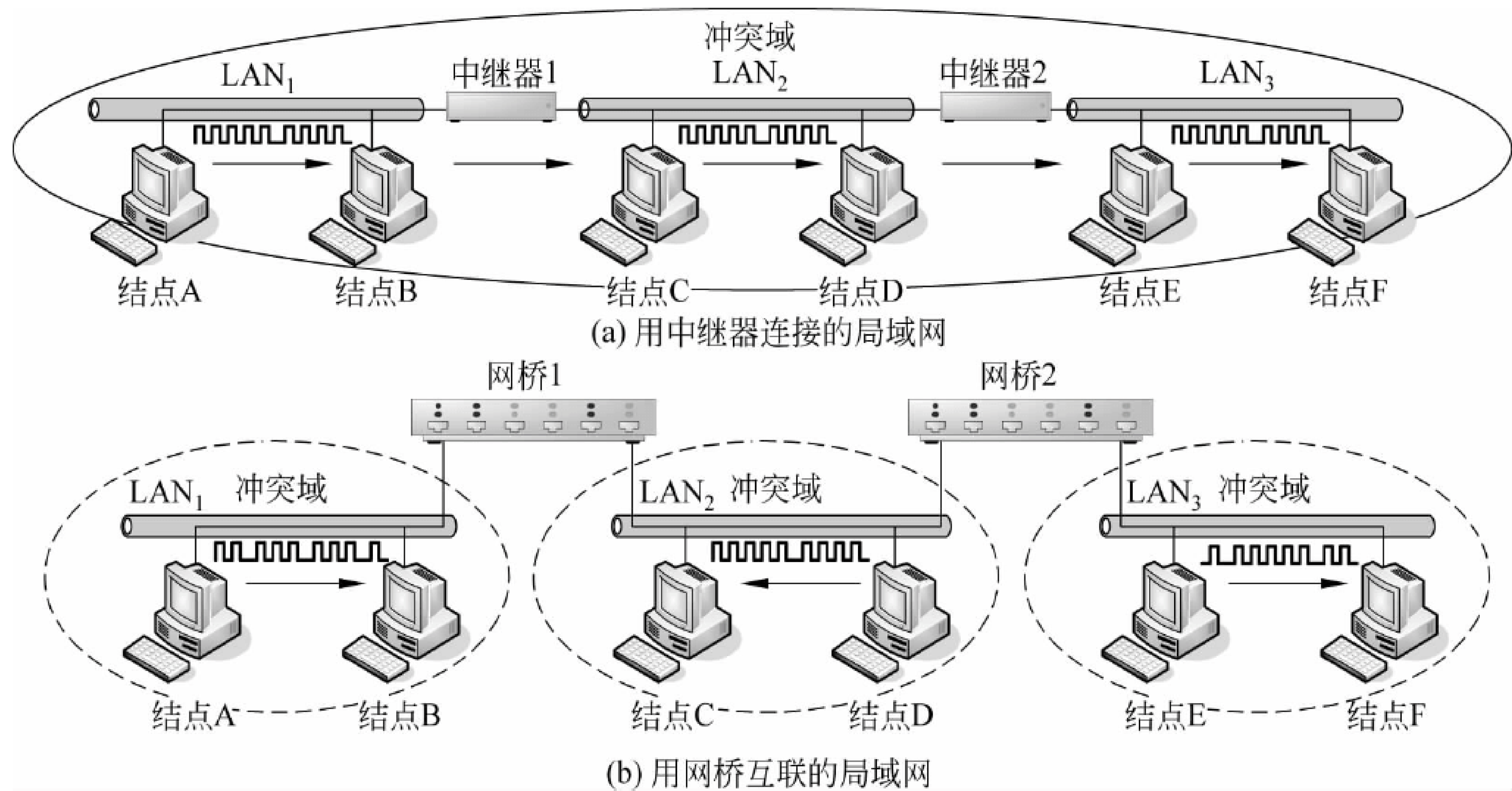


图 4-28 网桥与中继器作用的比较

(2) 当网桥接收到目的 MAC 地址为多播或广播地址,以及接收帧的目的 MAC 地址在转发表中查不出来,网桥无从决定应该从哪个端口转发时,它只能采用扩散算法,通过除进入的端口之外的所有端口转发出去,只要这个结点在互联的局域网中,那么广播的数据帧就有可能到达目的结点。这种方法很简单,但是却带来了很大的问题。这种盲目的广播会使网络中的无用通信量剧增,造成"广播风暴"。

2. 网桥与中继器、集线器、交换机的比较

在讨论了网桥基本工作原理之后,人们自然会想到网桥与中继器、集线器、交换机的区别问题。表 4-3 给出了中继器、集线器、网桥与交换机在协议层次、主要功能等方面的比较。

表 4-3 中继器、集线器、网桥与交换机的比较

比较内容	中 继 器	集 线 器	交 换 机	网 桥
协议层次	物理层	物理层	MAC 层	MAC 层
主要功能	连接 Ethernet 缆段,增加总线长度,增加接入的结点数量	接入多台计算机,形成星状结构的 Ethernet	连接多台计算机,实现快速帧转发	互联多个同构或异构的局域网
工作原理	信号放大与整形	信号放大与整形	在多端口之间建立并发连接与帧转发	MAC 地址过滤与帧转发
结构特点	两个端口	多端口	多端口	多为两个端口,也可以有多端口
使用地址			MAC 地址	MAC 地址
冲突域	连接在多个缆段上的所有结点属于一个冲突域	连接在集线器上的所有结点属于一个冲突域	端口独占,不存在冲突	每个互联的局域网分别是一个冲突域

总结:中继器、集线器、网桥与交换机的区别,可以从以下几个主要方面来看。

(1) 从网络协议层次的角度,中继器、集线器工作在物理层,而网桥与交换机工作在MAC层。

(2) 从使用局域网类型的角度,中继器、集线器是专为Ethernet设计,只是在Ethernet组网中才涉及的联网设备;而交换机可以有Ethernet交换机、Token Ring交换机等不同类型。

(3) 从设计目的的角度,中继器、集线器与交换机属于组建局域网需要使用的设备,而网桥属于在MAC层实现局域网互联的设备。

3. 路由器

路由器工作在网络层,使用IP地址进行路由选择。这个问题在网络层会详细讨论。

4. 网关

网关这一术语出现得很早。在ARPANET与TCP/IP开始研究的阶段,术语"网关"实际上不是后来人们常用的路由器。随着网络互联技术研究的深入,网关的概念更多用于网络层之上的传输层与应用层的异构网络互联的协议变换的设备。目前,在网络安全的防火墙技术中也会遇到"网关"的术语。在物联网体系结构中,必然会涉及3G/4G移动通信网、有线电话网、有线电视网与传统TCP/IP异构网络的互联,网关的概念与应用将会越来越多。

图4-29给出了交换机、网桥、路由器与主机在网络层次结构上的区别。

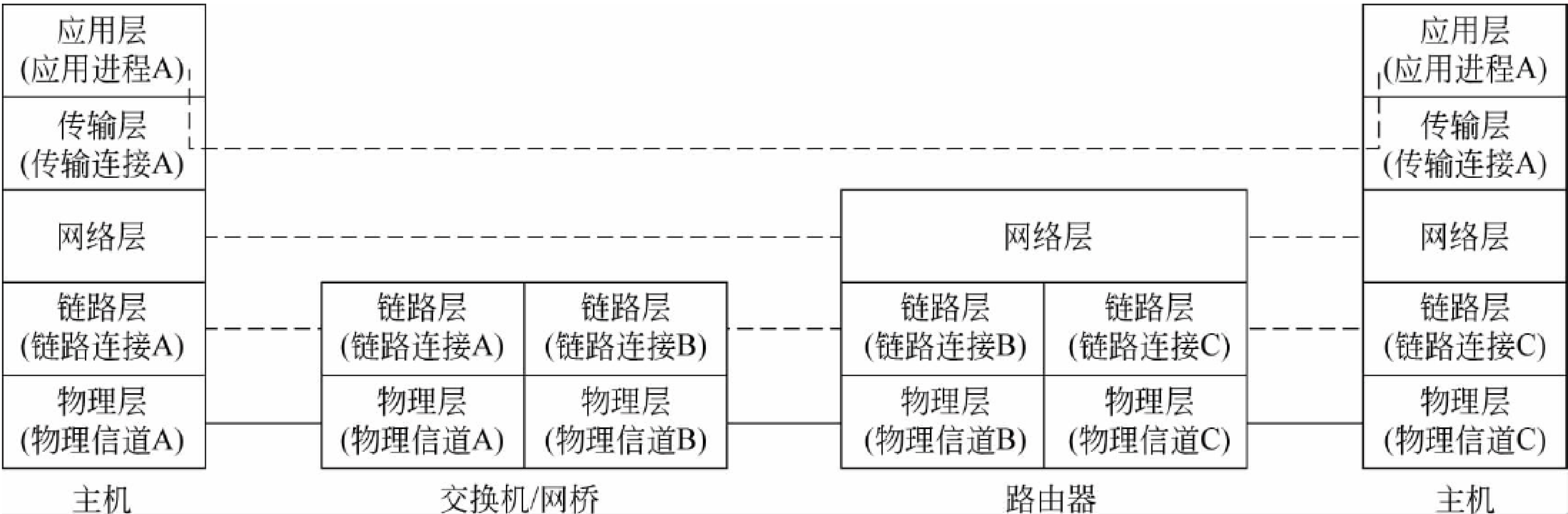


图4-29 从网络层次的角度认识主机与网络设备的区别

问题4-42:IEEE 802.11协议族是由哪些协议组成的?

1. IEEE 802.11协议族的组成

从1999年到2006年是无线网络,尤其是无线局域网发展最快的阶段,可以从表4-4中得出这样的结论。表4-4给出了按字母顺序列出的IEEE 802.11标准的主要内容与颁布时间。需要注意的是,表中同时列出了IEEE 802.16无线城域网的相关标准。

表4-4 无线网络标准

标准名称	标准描述	发布时间
IEEE 802.11a	传输标准,5GHz波段,速率54Mbps	1999年
IEEE 802.11b	传输标准,2.4GHz波段,速率11Mbps	1999年
IEEE 802.11d	多国漫游	2001年

续表

标准名称	标准描述	发布时间
IEEE 802.11e	服务质量 QoS	2005年
IEEE 802.11f	接入点第二层(数据链路层)漫游	2003年
IEEE 802.11g	传输标准,2.4GHz 波段,速率 54Mbps	2003年
IEEE 802.11h	动态频率选择与传输功率控制	2003年
IEEE 802.11i	无线通信安全	2004年
IEEE 802.11j	4.9～5GHz 波段传输标准	2004年
IEEE 802.11k	无线电资源管理	2002年
IEEE 802.11n	传输标准,5GHz 波段,速率 100Mbps	2006年
IEEE 802.11r	快速漫游	2006年
IEEE 802.11s	接入点无线网络	2006年
IEEE 802.11x	无线安全认证	2001年
IEEE 802.16d	高速无线城域网(固定应用)	2004年
IEEE 802.16e	高速无线城域网(移动应用)	2006年

2. 无线网络标准的分类

表4-4列出的无线网络标准可以分成以下三类。

1) 无线传输

IEEE 802.11a、802.11b 与 802.11g 分别定义了:5GHz 波段(速率 54Mbps)、2.4GHz 波段(速率 11Mbps)与 5GHz 波段(速率 100Mbps)等无线传输标准。

2) 无线网络安全

最早出现的无线网络安全标准是 IEEE 802.11x。IEEE 802.11i 用新的加密标准去替代静态密钥 WEP 标准,它可以实现密钥的动态分配。在 IEEE 802.11i 标准颁布之前,已有一个 WiMax 论坛的 WPA2 临时标准。制定 WPA2 临时标准的目的是防止无线网络厂商在正式标准出台之前,其产品与研究背离 IEEE 802.11i 的发展方向。

从协议数量和实际应用可以看出,无线网络安全研究还刚刚起步,是今后无线网络研究的一个重点领域。

3) 无线网络管理

无线网络管理除了包括传统网络管理的内容,还涉及对结点的漫游管理与保证无线网络传输服务质量 QoS 的问题。

IEEE 802.11e 为数据、语音与视频流量定义了无线网络传输能提供的服务质量 QoS 基本水平。

IEEE 802.11f 定义接入点之间在第二层(数据链路层)漫游时的通信,它支持在一个无线局域网内部的漫游,不支持在不同无线局域网之间的漫游,因为会涉及第三层的路由问题。

IEEE 802.11d 是 2001 年颁布的标准,它用于协调不同国家无线局域网之间的漫游问题。IEEE 802.11r 标准用来提高结点的不同无线局域网之间漫游的接入速度。

IEEE 802.11s 工作组是 2004 年成立的,主要研究如何将移动结点接入主干网的路由、负载平衡和性能优化问题。

IEEE 802.11h 用于动态频率选择与传输功率分配等物理层的问题。研究它的初衷是

准备在欧洲5GHz波段上，为接入点制定一系列的控制信号，以协调同一个波段的雷达与卫星信号，防止它们之间相互干扰。IEEE 802.11k是无线电资源管理标准。IEEE 802.11j是4.9～5GHz波段传输标准。

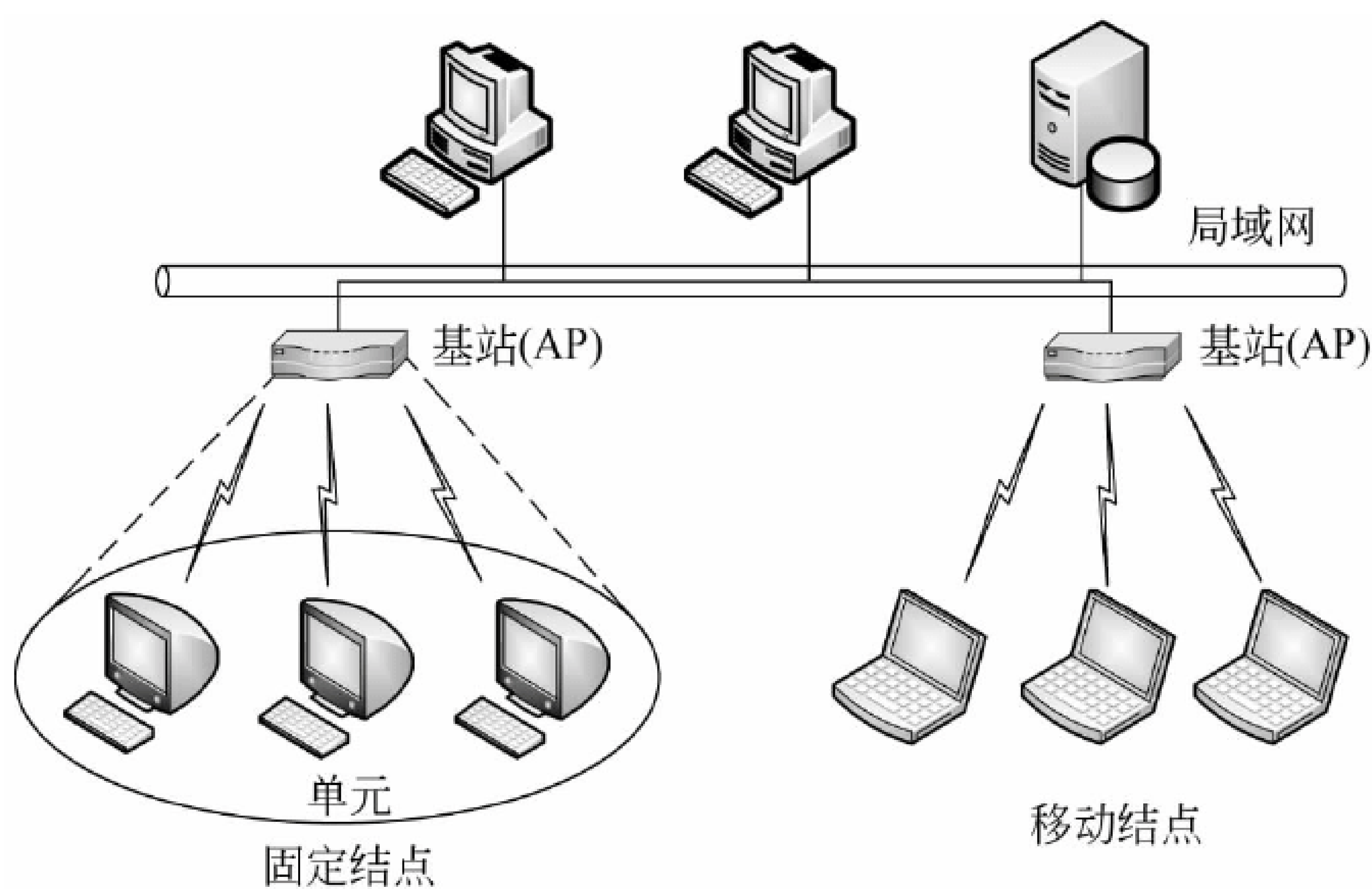
2006年5月，IEEE 802.11n标准在委员会的投票中以超过3/4多数通过，标志着业界对传输速率达到100Mbps的高速WLAN技术的高度关注。

问题4-43：什么是无线局域网中的“一跳”和“多跳”？

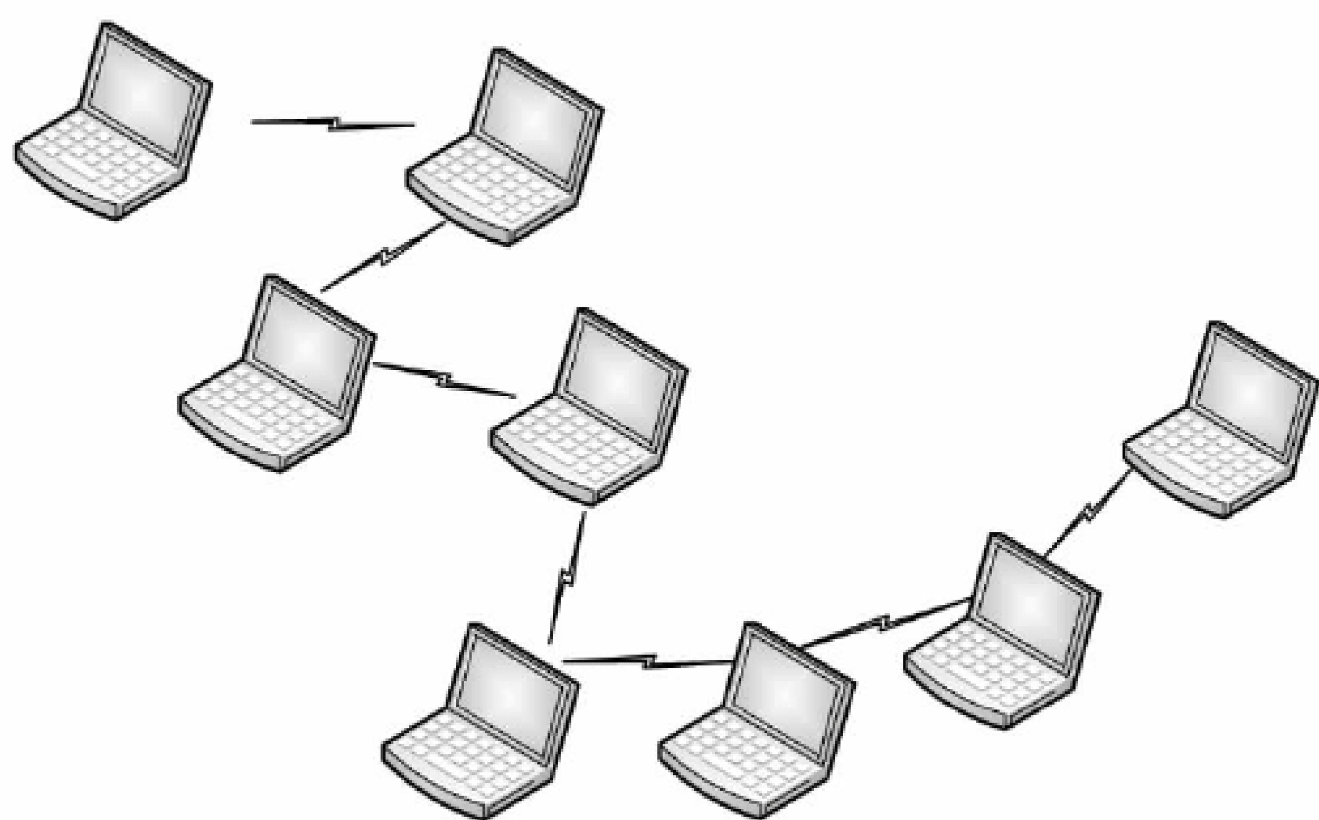
在讨论无线局域网MAC层工作原理时，经常会遇到术语“一跳”与“多跳”。“一跳”与“多跳”正好反映出有基站的无线局域网WLAN与无基站的无线自组网Ad Hoc的区别。图4-30给出了“一跳”“多跳”与WLAN、Ad Hoc的区别。

有基站的无线局域网WLAN的工作原理与Ethernet很相似。如图4-30(a)所示的固定或移动的无线局域网结点通过基站(AP)接入到无线局域网中。无线局域网中的结点只需要通过与基站AP之间的“一跳”传输，就可以通过AP实现数据交互。

如图4-30(b)所示的无线自组网Ad Hoc之间不存在基站AP，每个结点既是面向用户的端系统，同时又要起到路由器的作用，通过“多跳”转发来实现结点之间的数据交互。



(a) 有基站的无线局域网结构



(b) 无基站的Ad Hoc网络结构

图4-30　一跳与多跳示意图

问题4-44：如何理解“Wireless Fidelity”的含义？

在很多中文出版物或教材中将“Wi-Fi(Wireless Fidelity)”直译为“无线保真度”。如何

理解英文术语的确切含义,需要从技术内涵的角度去推敲。

1997年,IEEE公布了IEEE 802.11无线局域网标准。凡是读过无线局域网标准的技术人员都会感到标准涵盖的内容多,协议设计考虑到计算机网络与无线通信技术,链路层协议本身就很复杂。再加上无线通信技术发展迅速,物理层实现技术复杂,变化很快。在这样一个大的背景下,协议的制定者不可能在实现的技术细节上考虑得十分周全,因此不同厂商设计和生产的无线局域网产品一定会出现不兼容的问题。

针对这个问题,1999年8月由350家业界主要成员(如Cisco、Intel与Apple等)组成了致力于推广IEEE 802.11标准的Wi-Fi联盟(Wi-Fi Alliance)。其中,术语"Wi-Fi"或"WiFi(Wireless Fidelity)"涵盖着"无线兼容性认证"的含义。Wi-Fi联盟是一个非营利的组织,它授权在8个国家建立了14个独立的测试实验室,对不同厂商生产的802.11标准的无线局域网设备,如无线网卡、接入点AP、虚拟AP、无线网桥、无线路由器,以及采用802.11无线接口的笔记本、Pad、智能手机、相机、电视、RFID读写器进行互操作性测试,以解决不同厂商设备之间的兼容性问题。测试通过的网络设备都可以标记"Wi-Fi CERTIFIED"。因此,将"Wireless Fidelity"理解为"无线兼容性认证"似乎更为贴切。

问题4-45:术语辨析:BSS、ESS与MBSS。

2007年的IEEE 802.11标准定义了两类组网的结构模式:基础设施模式(Infrastructure Mode)与独立模式(Independent Mode)。

如果组建无线网络之前必须实现安装基站,Wi-Fi无线局域网中的基站是无线接入点AP,它相当于人们平时使用的手机的3G或4G基站一类的通信基础设施,那么它一定属于需要基础设施的结构模式。因此,基础设施模式也称为"基础结构型"。

基础设施模式可以进一步分为基本服务集(Basic Service Set,BSS)与扩展服务集(Extended Service Set,ESS)。

对应于独立模式的是独立基本服务集(Independent BSS)。独立基本服务集主要是指无线自组网(Ad Hoc)。

2011年的修正案IEEE 802.11s中又增加了第4种混合模式,对应的是Mesh基本服务集(MBSS)。

问题4-46:什么是分布式系统DS与无线分布式系统WDS?

为了扩大无线局域网的覆盖范围,可以通过Ethernet交换机将多个BSS互连起来,构成一个扩展服务集ESS,并可以通过路由器接入到Internet。ESS中的无线主机A可以通过基站AP1、Ethernet交换机、基站AP2与ESS中的任何一台无线主机通信;也可以通过基站AP1、Ethernet交换机与路由器接入主干网,访问Internet中的Web服务器或主机N,这样就构成了一个更大的分布式系统(Distribution System,DS)。分布式系统结构如图4-31所示。

如果通过无线网桥、无线路由器将多个BSS连接起来,就构成无线分布式系统(Wireless DS,WDS)。无线分布式系统结构如图4-32所示。分布式系统与无线分布式系统没有本质的区别。

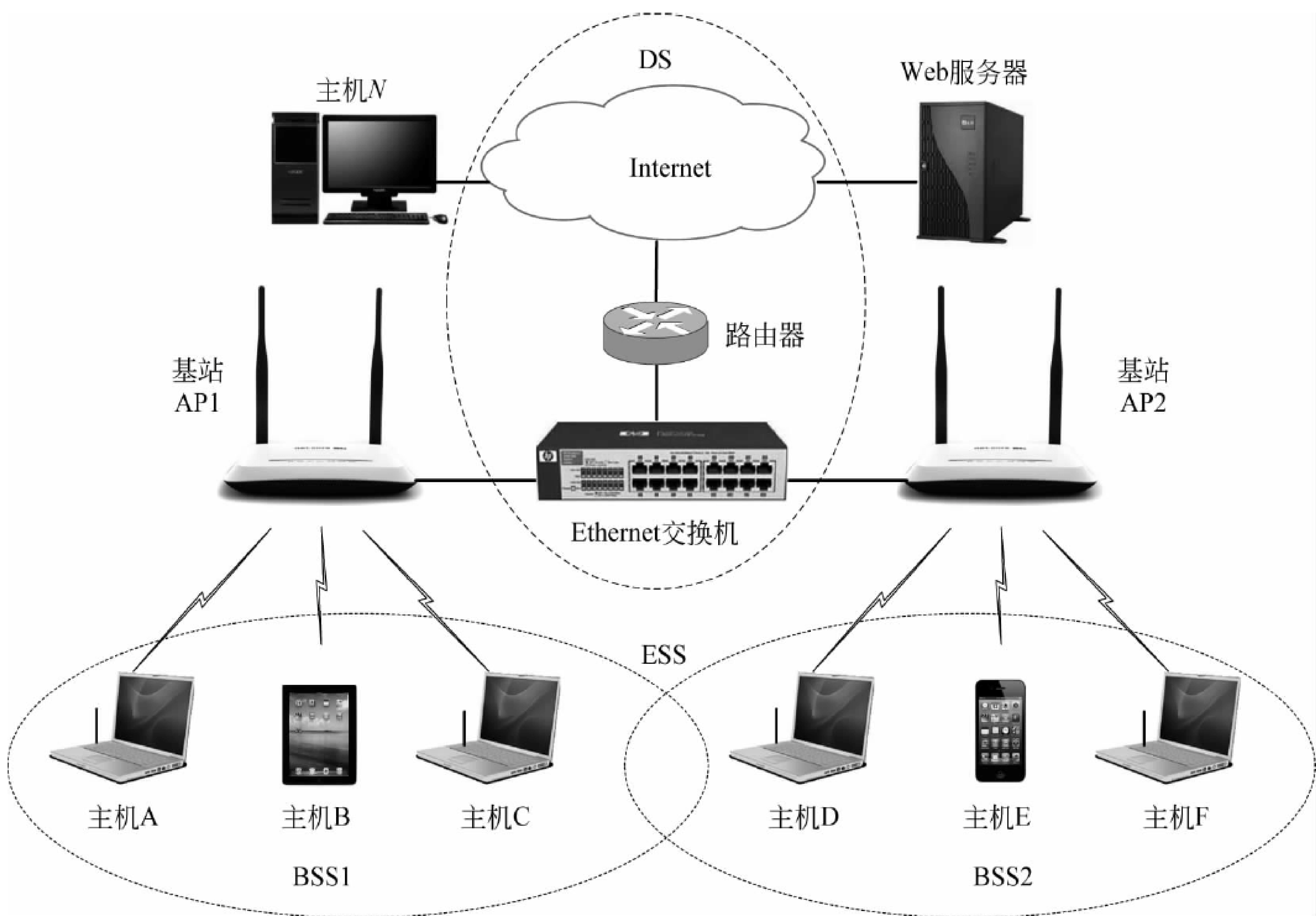


图 4-31　分布式系统结构示意图

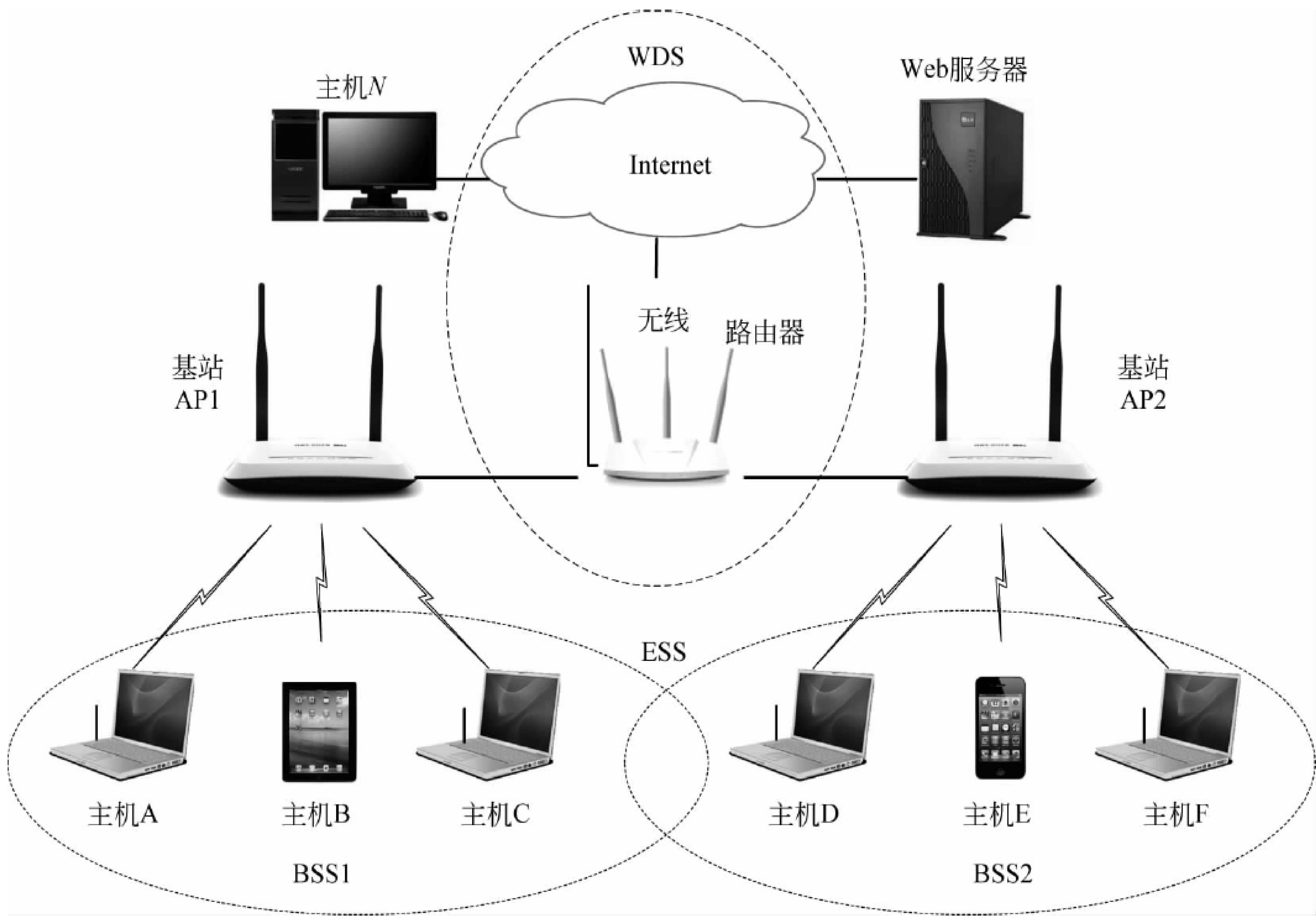


图 4-32　无线分布式系统结构示意图

问题 4-47：术语辨析：Ad Hoc 与 Mesh。

1. 无线自组网 Ad Hoc

无线自组网 Ad Hoc 的结构如图 4-33 所示。独立型无线自组网中没有无线基站，无线主机之间采用对等的点-点方式通信。不相邻无线主机之间的通信，需要通过相邻无线主机转接的多跳方式完成。

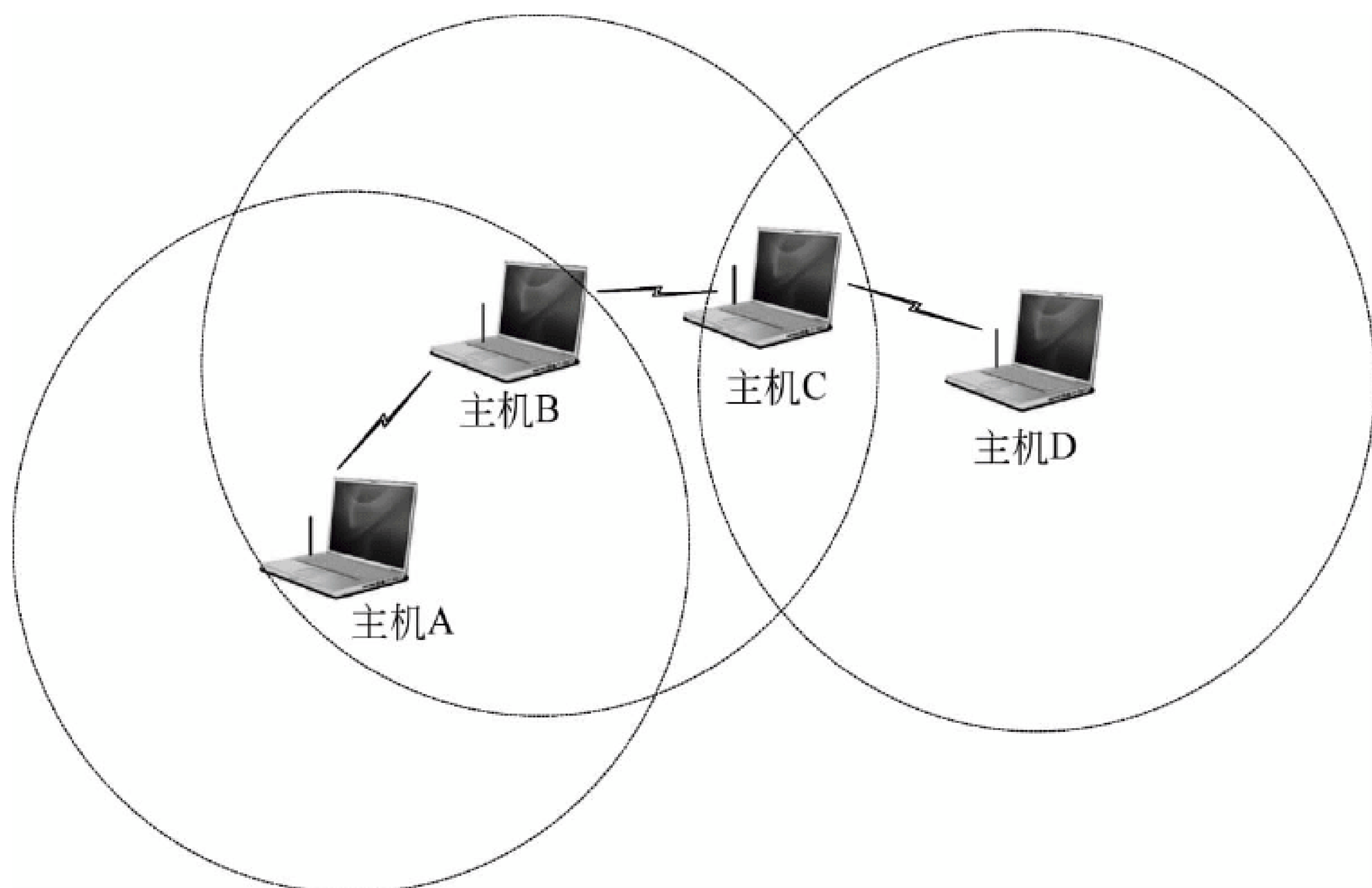


图 4-33　无线自组网中主机点-点通信方式示意图

Ad Hoc 网络具有以下几个主要的特点。

(1) 自组织与自修复。

(2) 无中心。

(3) 多跳路由。

(4) 动态拓扑。

由于 Ad Hoc 网络允许无线主机根据自己的需要开启或关闭，并且允许主机在任何时间、以任意速度、在任何方向上移动，同时受主机的接收信号灵敏度、天线覆盖的范围、主机的地理位置与主机之间障碍物遮挡，以及信号多径传输、信道之间干扰等因素的影响，使得主机之间的通信关系会不断变化，造成了 Ad Hoc 网络拓扑的动态改变。因此，要保证 Ad Hoc 网络的正常工作，就必须采取特殊的路由协议与算法。

2. 无线网状网

无线网状网(Wireless Mesh Network，WMN)的结构如图 4-34 所示。

无线 Mesh 网络是由一组呈网状分布的无线 AP 组成，AP 之间通过点-点无线信道连接，形成具有自组织、自修复等特点的“多跳”网络。

无线自组网 Ad Hoc 与无线网状网 WMN 的相同之处表现在：无线 Mesh AP 可以形成自己的 BSS，实现主机的接入功能，这一点与 BSS、ESS 相同；从“自组织”与“多跳”的角度，它与 Ad Hoc 网络相同，因此我们将无线 Mesh 网络归纳为混合型的网络。

无线自组网 Ad Hoc 与无线网状网 WMN 的不相同之处如下。

(1) 无线 Mesh 网络是通过 Mesh AP 与 Mesh AP 的点-点连接形成了网状网结构，而 Ad Hoc 网络直接由无线主机之间的点-点连接形成网状网。

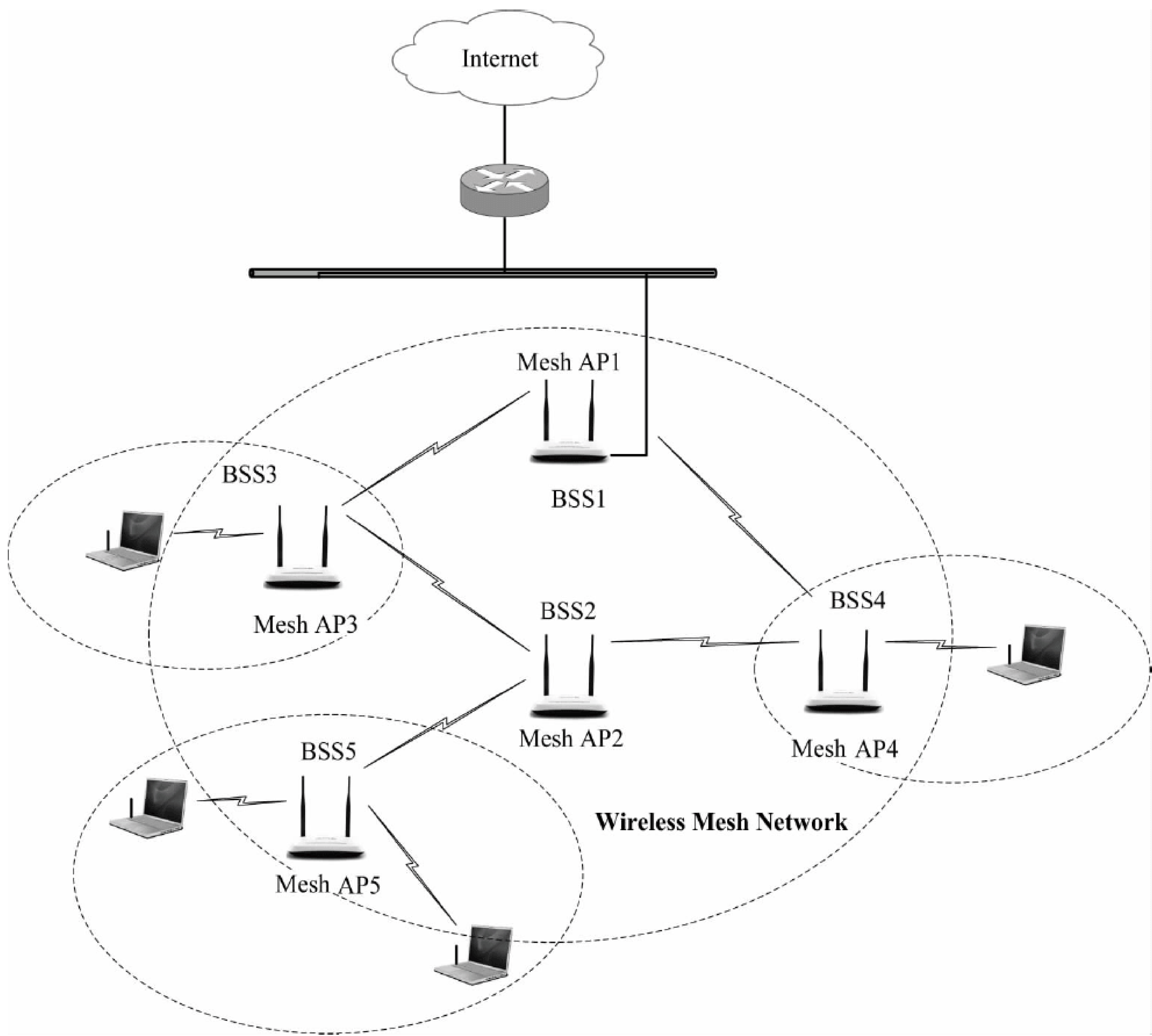


图 4-34　Mesh BSS 结构示意图

（2）无线 Mesh 网络主要适用于大面积、快速与灵活组网的应用需求；而 Ad Hoc 网络主要适用于多主机在移动状态下自主组网的应用需求。

问题 4-48：术语辨析：AP 与 Mesh AP。

从接入的角度，每个无线 AP 都可以形成自己的 BSS。无线接入点也可以作为无线网桥，通过无线信道在 MAC 层实现两个或两个以上的无线局域网，或无线局域网与有线局域网的无线桥接与中继的功能。

与传统的 AP 相比，由于无线 Mesh 网络中的 AP 增加了 MAC 层路由选择与自组织的功能，因此无线 Mesh 网络中的 AP 又叫作“Mesh AP”。AP 与 Mesh AP 的区别如图 4-35 所示。

问题 4-49：术语辨析：SSID 与 BSSID。

在 Wi-Fi 中必须解决 AP 设备与接入主机的识别问题。802.11 协议定义了 AP 的服务集标识符（Service Set Identifier，SSID）与基本服务集标识符 BSSID 的概念，两者息息相关但是又有很大的区别。

1．SSID

当网络管理员安装 AP 设备时，首先要为这个 AP 分配一个服务集标识符 SSID。按照 802.11 协议规定，AP 设备的名字最长为 32 个字符，并且区分字符的大小写。SSID 用来表

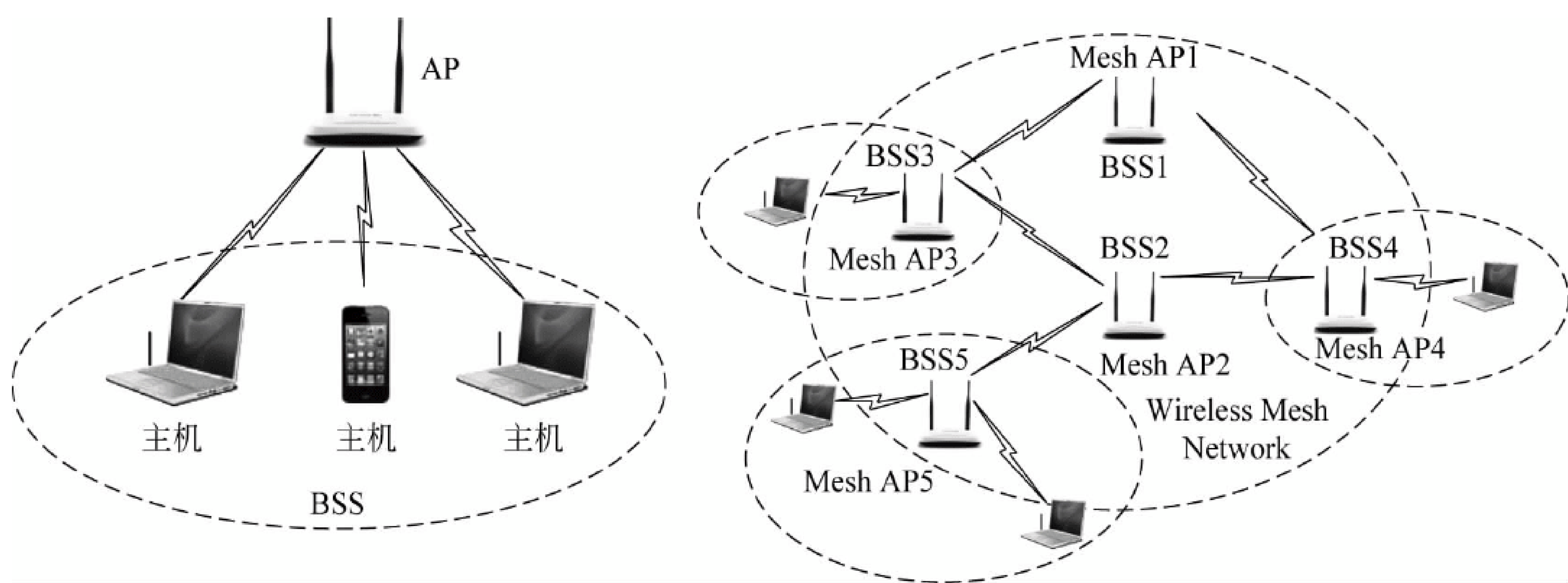


图 4-35　AP 与 Mesh AP 的区别

示以 AP 作为基站的 BSS 的逻辑名,它与 Windows 工作组名类似。例如,南开大学网络研究室教师办公室 AP1 的 SSID 名是“TP-NK-NETLAB”。那么,由这个 AP1 组成的 BSS1 的 SSID 名就是“TP-NK-NETLAB”。

2. BSSID

如果说 SSID 是 AP 的一层标识,那么基本服务集标识符(Basic SSID,BSSID)就是 AP 的二层标识。

理解 BSSID 的内涵需要注意以下三点。

(1) 接入点 AP 与无线主机之间的通信是通过内部的无线网卡实现的。大部分情况下,“基本服务集标识符 BSSID”就是无线网卡的 MAC 地址。为什么说是大部分情况下?因为有的网络设备生产商也允许使用虚拟 BSSID。

(2) IEEE 802.11 标准规定的无线网卡 BSSID 与 Ethernet 网卡的 MAC 地址很相似,长度都是 6 个字节(48 位)。不同之处在于:802.11 协议规定无线网卡 BSSID 的第一个字节的最低位为 0、倒数第二位为 1,其余的 46 位按照一定的算法随机产生,这样可以以很高的概率保证产生的 MAC 地址是唯一的。BSSID 作为 AP 设备唯一的二层标识,在无线主机的漫游中起到了重要的作用。

(3) SSID 是用户为 AP 配置的 BSS 无线局域网的逻辑名;BSSID 是网络设备生产商为 AP 配置的更精确的二层标识符。例如,AP1 的 SSID 为“TP-NK-NETLAB”,对应的 MAC 地址为“00:0C:25:60:A2:1D”。在无线局域网运行过程中,相对于 Ethernet 的 MAC 地址。

问题 4-50:术语辨析:点协调、分布式协调与混合协调。

IEEE 802.11 的 MAC 层协议支持两种基本的访问控制方式:点协调功能与分布协调功能。理解它们的区别,需要注意以下几个问题。

(1) 点协调功能(Point Coordination Function,PCF)属于无争用服务。无争用服务系统的中心是基站——无线接入点 AP。在点协调功能工作模式中,基站 AP 控制着多个无线主机对共享无线信道的无冲突访问,形成了以基站为中心的星状网络结构,因此点协调功能 PCF 模式提供的是无争用服务。

(2) 分布协调功能(Distributed Coordination Function,DCF)属于争用服务。IEEE

802.11 的 MAC 层也可以采用载波侦听多路访问(CSMA)/冲突避免(Collision Avoidance,CA)的介质访问控制方法。

(3) IEEE 802.11 标准规定 MAC 层都必须支持分布协调功能 DCF,而点协调功能 PCF 是可选的。在默认状态下,802.11 的 MAC 层工作在分布协调功能 DCF 模式;只有在对传输时间要求高的视频、音频会话类应用时,才会启用点协调功能 PCF。

(4) 有些应用需要 Wi-Fi 提供比 DCF“尽力而为”服务更高一级的服务,但是又不需要 PCF 集中控制的服务,于是开始研究混合协调(Hybrid Coordination Function,HCF)控制方式,但是目前 HCF 控制方式仍处于研究阶段,没有相应的协议标准。

问题 4-51:为什么无线局域网不能采用 CSMA/CD 介质访问控制方法?

为了说明这个问题,首先要搞清 Ethernet 能够使用 CSMA/CD 介质访问控制方法的前提条件。回顾 CSMA/CD 介质访问控制方法讨论时有以下两个前提条件。

(1) 一个结点发送帧时要在“冲突窗口”之内,可以检测出发送帧是否出现冲突。IEEE 802.3 协议将 10Mbps 的冲突窗口定为 51.2μs,将帧的最短长度定为 64B。

(2) Ethernet 网卡能够用 Manchester 编码违例或根据接收数据流时钟的方法发现总线空闲和是否出现发送冲突。

Wi-Fi 在无线信道上无法满足这两个前提条件。原因有以下两个。

(1) 无线网卡的发送功率与接收功率一般相差都非常大。要求无线网卡在发送信号的同时,要处理微弱的接收信号,并判断是否出现冲突,从电路实现的角度难度很大,即使可以实现但成本很高。

(2) 无线通信环境中信号传输路径复杂。发送端的无线信号可能是经过绕射、折射、反射的多路径到达接收端,不能简单地根据不同主机之间的直线距离去估算传输延时和“冲突窗口”的数值。

因此,传统 Ethernet 与 Wi-Fi 无线局域网在信道访问控制方法上有以下两点不同。

(1) 传统 Ethernet 的结点在监测到总线空闲时,立即发送帧;而 Wi-Fi 无线局域网结点在监测到无线信道“闲”时,不是“立即”发送帧,而是要求所有准备发送数据帧的主机都执行退避算法,通过“冲突避免”来减小冲突发生的概率。

(2) 传统 Ethernet 发送结点只要在“冲突窗口”的时间内没有检测出冲突,就确定为发送成功,不需要接收结点发送确认帧;而 Wi-Fi 无线局域网发送结点需要等待接收结点发送回的确认帧,来判断此次发送是否成功。

问题 4-52:如何理解 802.11 协议对“漫游”的处理方法?

我们一般理解:漫游是指无线主机在不中断通信的前提下,在不同 AP 覆盖范围之间移动的过程。Wi-Fi 必然要支持无线主机的漫游。从 MAC 层来看,漫游是无线主机转换 AP 的过程。从网络层及以上高层来看,漫游是在转换接入点的同时仍然维持原有的网络连接的过程。

但是,IEEE 802.11 标准中并没有用到“漫游”这个术语。人们对这种现象的解释是:“不论何时何地,漫游都是客户端的自由”。IEEE 802.11 将是否支持漫游的问题交给了无线局域网的网络软硬件厂商去自行决定。

在设计无线网络拓扑时,一定要考虑到无线主机无缝漫游的问题。IEEE 802.11 协议

设计了一套完整的管理帧与对无线主机在ESS中自由切换AP的管理机制。

无线网卡和AP设备有两种基本设计思路:一旦关联到一个AP之后就一直坚持着,直到完全接收的信号质量很差时才考虑转换接入点AP;一旦找到新的信号最强的AP就立即转换。AP通过周期性广播“Beacon”帧;无线主机通过被动扫描或主动扫描的方式来发现AP;通过发送“重关联请求帧”来启动漫游的过程。

问题4-53:802.11协议是如何支持移动终端设备节能管理的?

理解IEEE 802.11对移动终端设备节能管理的支持,需要注意以下几点。

(1)由于接入无线网络中有大量的设备是笔记本、智能手机与各种手持移动终端,因此节能非常重要,它关系到终端的移动性与续航能力。IEEE 802.11协议在帧控制字段中设置了一位“电源管理”位。

(2)IEEE 802.11支持两种电源管理模式:主动模式(Active Mode)与节电模式(Power Save Mode)。由于处于主动模式的数据传输速率高于节能模式,因此办公室的无线主机由于一直可以连接220V电源上,因此一般都处于默认的主动模式状态。IEEE 802.11协议默认的是主动模式。主动模式表示网卡处于时刻准备发送或接收数据的状态。

(3)很多移动终端设备是由内部电池供电,为了延长设备使用时间,可以选择为节能模式。节能模式是可选的模式。在节能模式中,主机要关闭无线发射与接收电路,处于“休眠”状态。协议规定:电源管理位为0,表示源主机在发送完该帧后仍然处于工作状态;电源管理位为1,表示源主机在发送完该帧后进入休眠状态。

节能模式对于移动互联网与物联网终端设备的研发是十分重要的。

问题4-54:如何理解AP的“双频多模”?

随着IEEE 802.11协议标准的不断完善,“双频多模”成为无线接入点AP研发与应用的重要方向,它可以适应多种工作环境,最大限度地发挥Wi-Fi的优势与特点,有效地解决无线主机的无缝漫游问题。

IEEE 802.11a、802.11b与802.11g等物理层标准的不同,导致了不同标准的无线设备之间存在着兼容性问题。IEEE 802.11a工作在5GHz,而IEEE 802.11b、802.11g工作在2.4GHz;IEEE 802.11a与802.11b发送信号所采用的调制方式也不相同。那么,一台无线主机漫游到不同物理层标准的BSS区域时就必须使用不同的无线网卡,这显然是不合适的。为了解决这个问题,无线AP设备的研制向着“双频多模”方向发展。其中,“双频”是指可以支持2.4GHz与5GHz两种频率;“多模”是指可以自动识别和支持IEEE 802.11a、802.11b与802.11g等多种物理层标准。图4-36给出了一种典型的“双频双模”(IEEE 802.11a与802.11g)AP的结构示意图。

问题4-55:如何认识“统一无线网络”研究的必要性?

第一代无线接入点AP相当于Ethernet Hub。AP设备通过无线信道与一组无线主机关联,作为BSS的中心结点执行CSMA/CA的MAC算法,实现无线主机之间通信的功能。

第二代无线接入点AP将无线接入与无线局域网管理功能结合到Ethernet交换机中,构成了ESS无线网络。

第三代无线接入点AP与无线局域网控制器结合,构建更大规模、集中管理的统一无线网络系统。

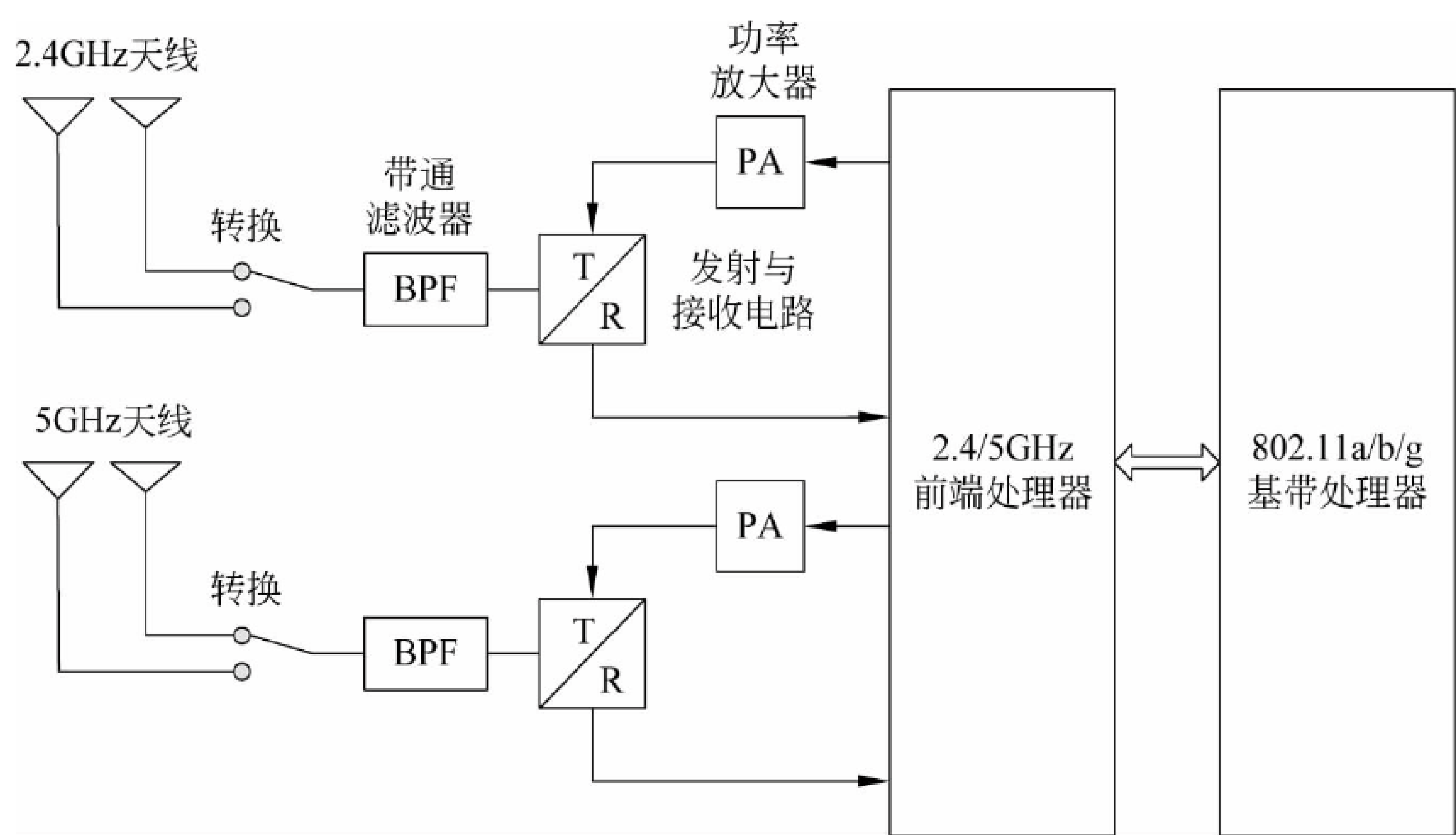


图 4-36　双频双模 AP 的结构示意图

理解更大规模、集中管理的“统一无线网络系统”技术研究的必要性，需要注意以下几个问题。

1．统一无线网络的基本概念

随着 Wi-Fi 从初期的家庭、小型办公室环境的应用，不断扩大到覆盖一个校园、一家大型医院或一个科技园区，从几个 AP 设备扩展为由数百个 AP 设备的大型无线网络系统，促使 Wi-Fi 网络结构从初期以自主 AP 为中心的基本服务集 BSS，发展到用 Ethernet 交换机将多个 BSS 互联起来构成的扩展服务集 ESS，直到将 Ethernet 交换机变换为无线局域网控制器（Wireless LAN Controller，WLC），出现了集中管理的大型无线网络结构。Cisco 将这种集中管理的无线网络结构命名为“Cisco 统一无线网络（Cisco Unified Wireless Network，CUWN）”。CUWN 的中心是无线局域网控制器 WLC。目前，CUWN 的概念已经被很多无线网络设备制造商所接受。典型的 WLC 集中式管理的无线网络结构如图 4-37 所示。

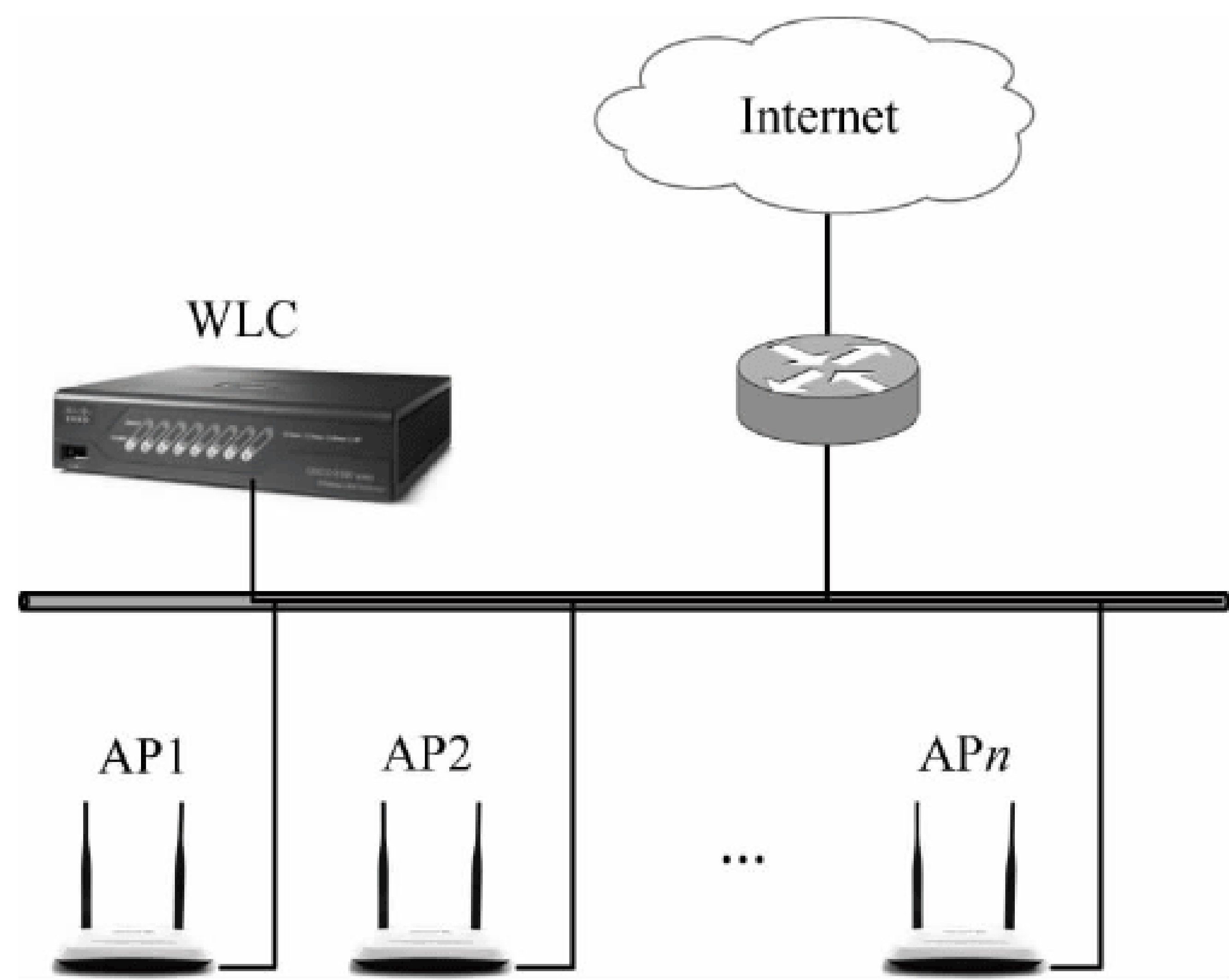


图 4-37　典型的 WLC 集中式管理的无线网络结构

2．统一无线网络的特点

推动 Wi-Fi 结构由自治方式到集中方式的转型的动力主要来自大型无线网络运行、维护与网络管理的压力。集中式管理的统一无线网络的特点主要表现在以下几个方面。

(1) 自治 AP 设备中有很多参数需要配置。在一个大型的 ESS 系统中,为了简化配置与维护的工作量,网络管理员一般是将所有 AP 的参数配置成相同的值。即便是这样,网络管理员也要实地对每一个 AP 设备进行设置。而在集中管理的统一无线网络系统中,网络管理员可以通过 WLC 的控制界面,在很短的时间内完成所有 AP 参数的配置。对于同样规模的无线网络,更新和修改多个自治 AP 的参数配置可能需要几个小时甚至是几天的时间,而对于 WLC 来说只需要几秒钟。在统一无线网络中增加一个新的 AP,它能够根据 WLC 已经定义的参数进行自我配置。

(2) 在实际运行的系统中,很难保证所有自治 AP 运行相同版本的软件,网络管理员需要为每一个 AP 单独更新现有版本的升级软件、缺陷修补补丁,添加新的功能。而集中式管理的统一无线网络系统中,所有的 AP 运行着相同软件的镜像。网络管理人员可以方便地为所有 AP 更新软件。

(3) 在设计一个大型的无线网络系统结构时,网络技术人员需要实地勘察无线网络工作的环境、覆盖范围与用户数量,以确定 AP 的数量与位置,并且需要从减少干扰的角度完成 AP 信道复用的规划,为不同位置的 AP 配置不同的发射功率。这就需要网络技术人员有很好的无线通信技术知识与无线网络安装、配置、运维的经验。在日常运行过程中,网络技术人员需要根据外部环境的变化(建筑物内新增墙体、设备或家具),建筑物中用户人数的变化,以确定 AP 设备数量的增减;需要根据周边环境中出现新的干扰信号,如无线局域网、蓝牙设备、微波炉或视频设备产生的相同或相近频率的信号产生干扰,来决定 AP 安装位置的变化,或者需要选择新的信道频率、改变信号功率,以保证网络系统的正常运行。很多移动计算应用都需要无线网络系统保证无线主机的无缝漫游。自治 AP 系统解决的方法只能是不断地通过人工方式去调整和部署冗余的基础设施,增大 BSS 之间重叠的面积。完成以上网络系统维护任务的工作量很大,需要使用无线测量设备,并且对网络管理人员的技术水平要求也很高。

(4) 为了解决这些问题,统一无线网络增加了"无线资源管理(Radio Resource Management,RRM)"功能。RRM 又称为"Auto-RF"。无线资源管理通过连续地采集和监测来自多个 AP 无线信道的数据,利用无线资源管理算法,分析无线通信系统的状态,通过协调多个 AP 的信道频率与功率设置,来提高信号传输质量,增强对无缝漫游的支持能力。Auto-RF 可以降低无线网络系统的维护难度,提高了无线网络运行的可靠性与可用性。

问题 4-56:什么是胖 AP 与瘦 AP?

在出现统一无线网络 UWN 的概念之后,人们将不使用无线局域网控制器 WLC 的 AP 称为"自治"或"基于 IOS(Internetwork Operating System)的 AP"。所谓"自治"是指:传统的无线接入点 AP 的操作系统与配置文件存储在设备的存储器中,可以作为一个完整的系统独立地工作。自治 AP 系统的功能是通过两类进程(实时进程与管理进程)来实现。实时进程主要包括:无线信号的发送与接收、MAC 协议工作过程的控制与管理、加密;管理进程主要包括:无线信道频率与发射功率的管理、关联与漫游的管理、客户端认证、安全与 QoS 管理。

在统一无线网络中,WLC 按照无线接入点控制与配置(Control And Provisioning of Wireless Access Point,CAPWAP)协议,对大量 AP 的管理进程实现集中管理。因此,人们将统一无线网络中的 AP 称为"瘦 AP"或"轻量级接入点 LAP",将自治 AP 称为"胖 AP"或

“分离 MAC 架构”。胖 AP 与瘦 AP 在功能上的区别如图 4-38 所示。

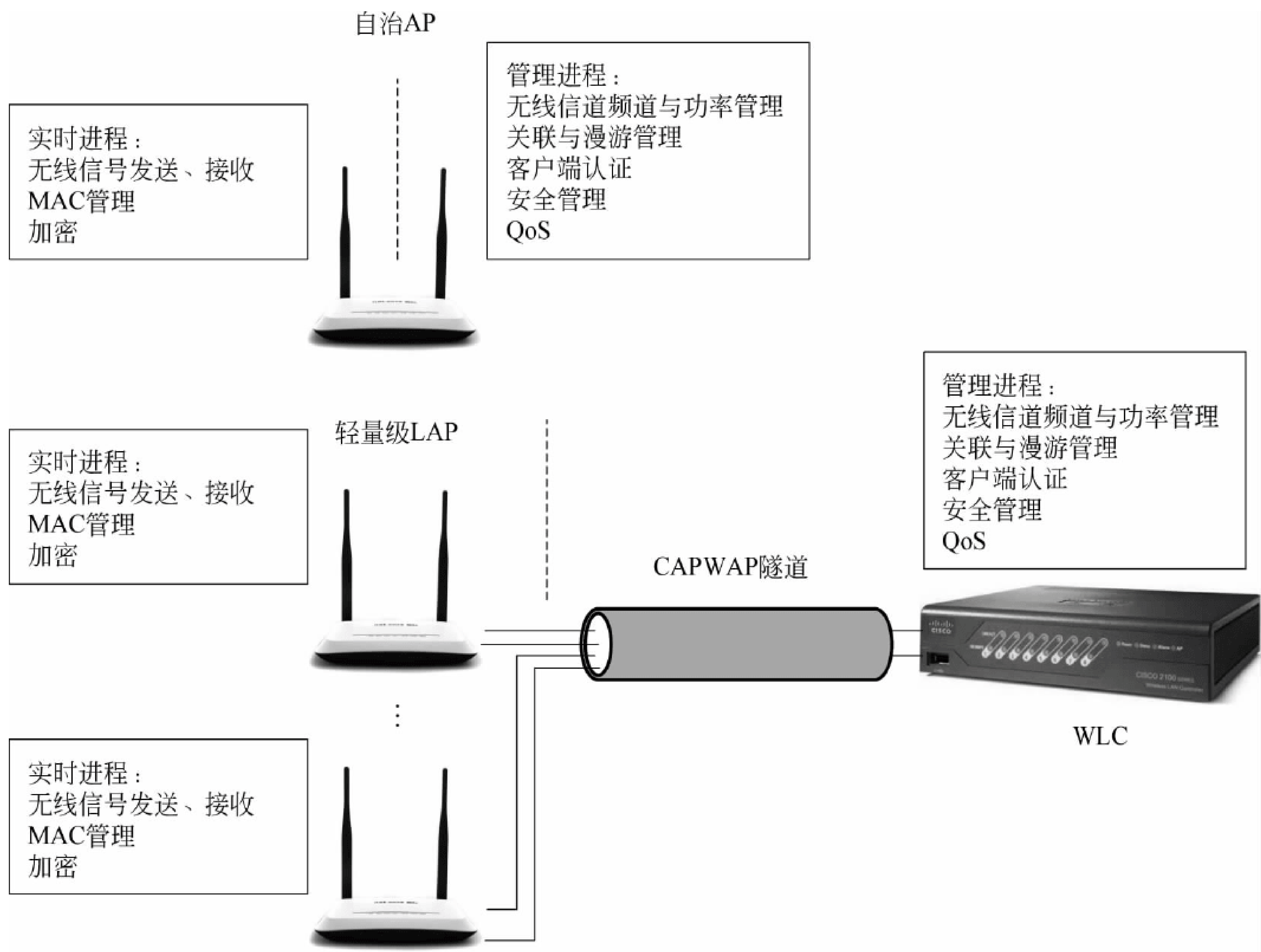


图 4-38 胖 AP 与瘦 AP 在功能上的区别

问题 4-57：如何理解虚拟 AP?

理解虚拟 AP 的概念，需要注意以下几个问题。

1. 传统 AP 应用的局限

早期的机场考虑到乘客上网、在线购物与支付的需求，为乘客接入 Internet 组建了专门的 Wi-Fi 网络，而很多其他的应用(如登机口的航空检票设施、零售柜台)也需要使用 Wi-Fi 接入，解决的办法只能是另建一个 Wi-Fi 网络。因此，传统的方法不能在一个 AP 构成的 BSS 中为不同类型的用户提供区分服务，需要分别构建和管理多个物理网络。这种解决方案带来了无线网络建设上的重复投资，增加了网络管理人员的维护工作量与成本，以及 AP 设备的位置、供电、无线频率配置困难的问题。针对这个问题，研究人员研发了用一个(组)物理网络基础设施去构建多重逻辑网络的虚拟接入点(Virtual AP)技术。

2. 虚拟 AP 概念的提出

用虚拟接入点方法构建的无线网络逻辑结构如图 4-39 所示。虚拟 AP 技术允许网络管理员在一组 AP 设备上，设置和控制多个动态 VLAN。从图中可以看出，该网络设置了三个虚拟网络。其中，Network A(SSID1)是一个公司的内部无线网络。如果用户要访问该网络，必须在公司网络的 Radius Server 上有账户。Network B(SSID2)是一家无线互联网接入服务提供商(WISP)。WISP 使用基于 Web 的身份认证系统，为注册的合法用户提供 Internet 接入服务。Network C(SSID3)用于提供 IP 语音服务，并且配备了 IP 用户级电话

交换机(Private Branch Exchange,PBX)。虚拟AP分别给对应SSID1的Network A、对应SSID2的NetworkB、对应SSID3的Network C分别分配一个虚拟的MAC地址BSSID A、BSSID B与BSSID C。

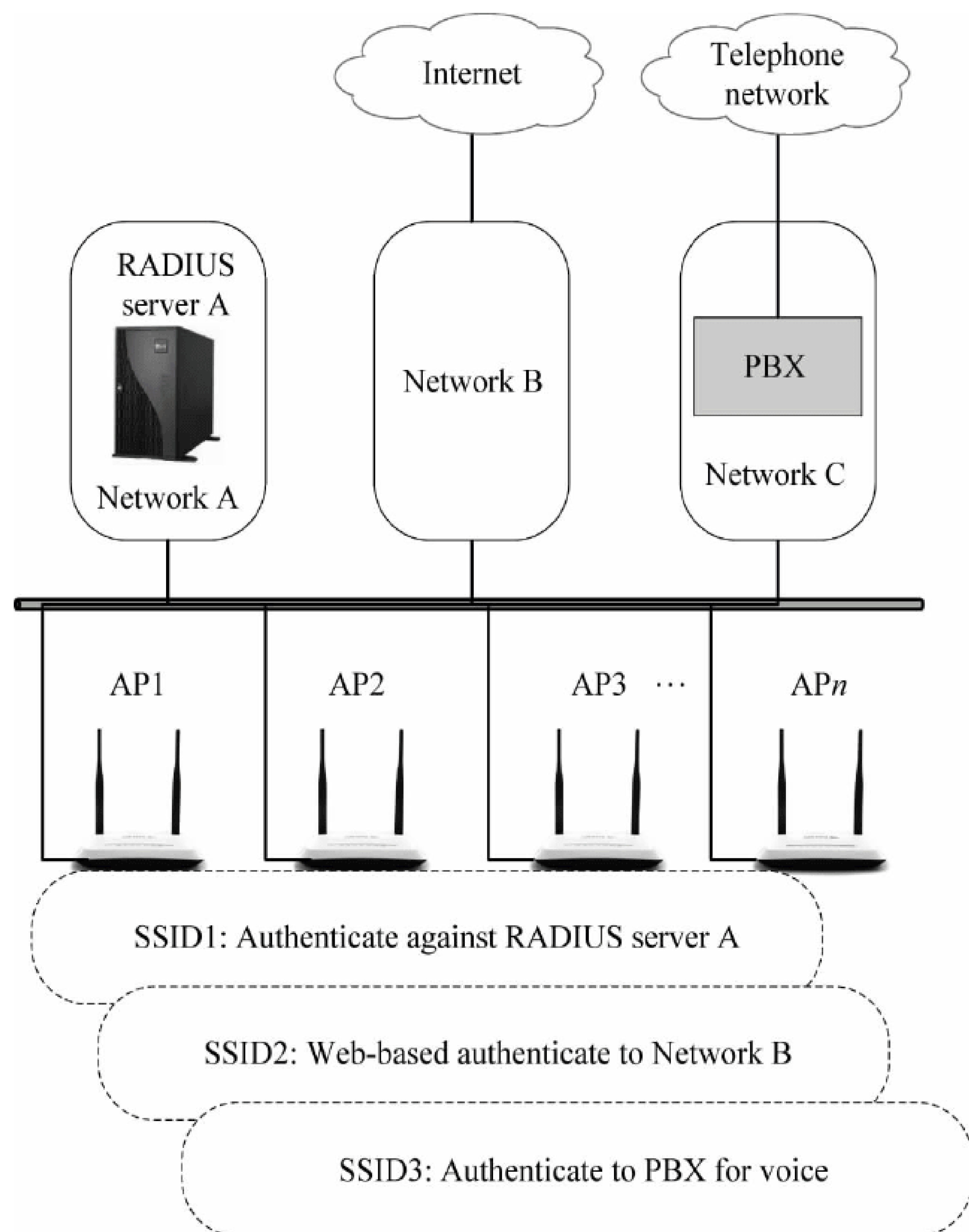


图4-39 虚拟接入点组网的逻辑结构示意图

由于所有AP可以使用虚拟的MAC地址BSSID A、BSSID B与BSSID C广播信标帧;无线主机可以接入到任何一台AP上,访问Network A、Network B与Network C。对于Network A的用户来说,只知道他的无线主机接入到服务集标识符为SSID1、MAC地址为BSSID A的公司AP中,不需要知道具体接入在哪个AP上,也不需要知道他实际上可能在多个物理的AP接入点之间漫游。

由于在共享无线基础设施的前提下,虚拟接入点组网方案可以为不同类型的用户提供区别服务,而共享的无线基础设施是由一个机构建设和管理,因此这种组网方案既节约了建设资金,避免了重复投资,又可以免于频率之争,便于统一管理与运营。由于虚拟接入点组网方案具有以上的优点,因此引起了人们越来越多的关注。

第三部分 习题参考答案

1. 最小帧长度为100b。
2. 如果最小帧长度减少800b,那么总线两端最远两台主机之间的距离至少为80m。
3. (1) 主机A检测到冲突需要的时间为10μs。

（2）当检测到冲突的时候，主机A已经发送数据1000b。

4.（1）从开始发送数据到检测到冲突，最短需要的时间是：$2000/2\times10^{8}=10\mu s$，最长需要的时间是：$20\mu s$。

（2）主机A有效传输速率约为9.33Mbps。

5. B1、B2转发表如表4-5和表4-6所示。

表4-5　B1转发表

目的地址	端口
H1	1
H5	2
H3	2
H2	1
H6	1
H4	2

表4-6　B2转发表

目的地址	端口
H1	1
H5	2
H3	1
H2	1
H6	2
H4	1

6. 根据2.4GHz信道复用规划方法，填充的信道如图4-40所示。

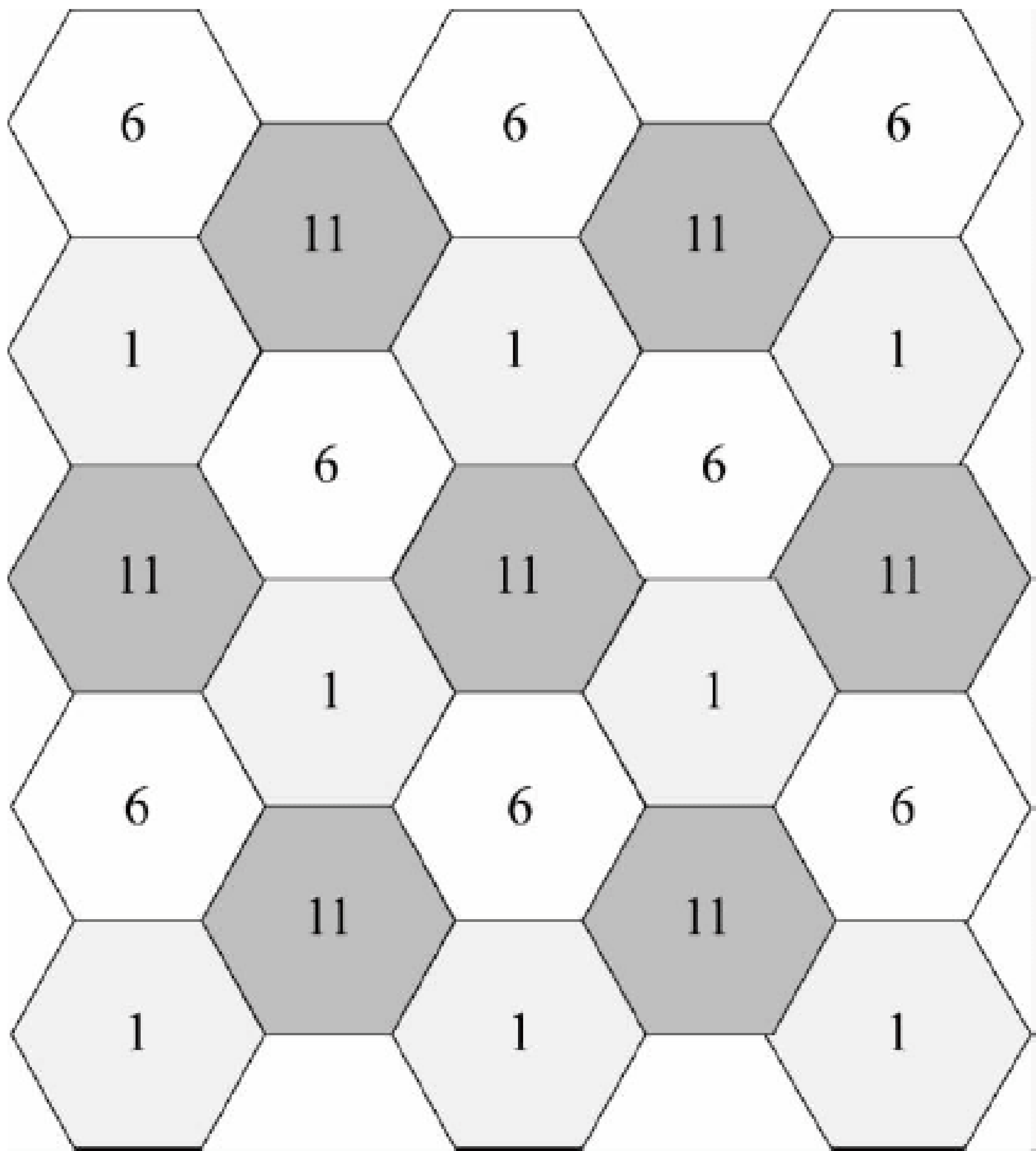


图　4-40

7. 管理帧中的关联请求帧

管理帧中的信标帧

控制帧中的CTS帧

控制帧中的ACK帧

数据帧

8. ① 0

② 1

③ 目的地址：00.16.11.22.b8.60

④ AP地址：00.57.55.66.aa.11

⑤ 源地址：01.a6.6e.00.18.a0

第 5 章 网络层

第一部分　学习目的、要求与知识点结构

1. 学习目的

本章在物理层与数据链路层讨论的基础上，系统地介绍网络层的基本概念、网络层向传输层提供的服务功能、路由选择算法与协议、流量控制算法、IP 协议，以及网络互联概念方法、路由器工作原理与设计方法。

本章的内容围绕着 IP 协议展开，通过本章的学习为进一步研究 Internet 工作原理与实现技术打下坚实的基础。

2. 学习要求

（1）理解：网络层与网络互联的基本概念。

（2）掌握：IPv4 协议的基本内容。

（3）掌握：IP 地址、路由算法与路由协议的基本概念。

（4）掌握：地址解析 ARP 的基本概念与方法。

（5）掌握：路由器与第三层交换的基本概念。

（6）掌握：MPLS 协议与 MPLS VPN 的基本概念。

（7）掌握：IPv6 协议的基本内容。

（8）掌握：移动 IP 的基本概念。

（9）掌握：ICMP 与 IGMP 的基本概念。

3. 本章知识点的组织与结构

网络层是计算机网络课程学习的重点，对于理解 Internet 实现技术至关重要。掌握这一部分知识对提高学生实际工作技能和今后继续学习的能力是十分重要的，同时本章的教学也是整个网络课程的重点与难点。为了帮助教师做好教学工作，本章的教师用书侧重于概念的讨论，习题指导书将给出各类帮助学生掌握知识的例题与习题。

图 5-1 给出了本章知识点结构。

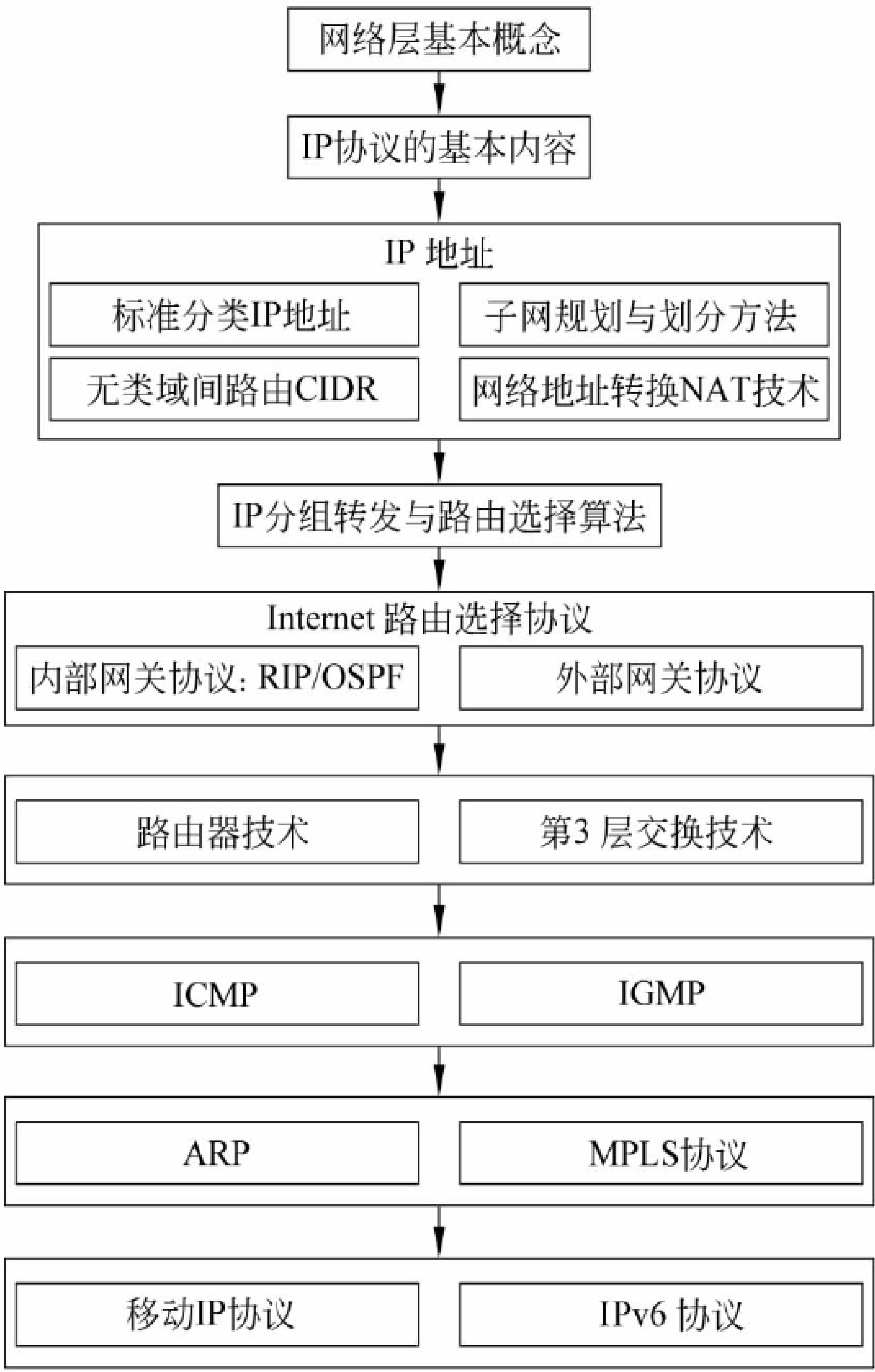


图 5-1　知识点的组织与结构

第二部分　教学内容问答

问题 5-1：如何评价 IPv4 协议？

如果回顾一下 IP 协议研究的时间、初期 IP 协议讨论的内容、IP 协议整个改进和完善的过程，以及实际取得的效果，可以从中得出的启示是：IPv4 协议的设计是成功的。它的成功表现在以下两个方面。

（1）实际的技术路线是正确的。如果要求计算机科学家在 20 世纪 70 年代就能够遇见计算机网络可能形成的规模、应用与问题，那不符合规律。IP 协议的设计者在第一个设计文档中只对 IP 分组结构做出了规定，对 IP 地址按照标准分类的方法给出了意见，提出了直接交付与间接交付、路由选择的概念，用简单的方法解决复杂问题，用“尽力而为”的服务去应对互联网络中存在的各种复杂的问题。这样做才有利于技术的推广与应用，才在 TCP/IP 体系与 OSI RM 的竞争中赢得了时间与市场。

（2）伴随着 Internet 规模的扩大和应用的深入，作为 Internet 核心协议之一的 IPv4 协议也一直处于一个不断补充、完善和提高的过程，但是 IPv4 版本的主要内容没有发生任何实质性的变化。实践证明，IPv4 是健壮和易于实现的，并且具有很好的互操作性。它本身也经受住了 Internet 从小型的科学研究范围中应用的互联网络，发展到今天这样的全球性

大规模网际网的考验,这些都说明IP协议是成功的。

问题5-2:如何理解IP协议的“尽力而为”服务的含义?

理解这个问题需要注意以下两点。

(1) 在讨论IP协议特点时经常会说:IP协议的“尽力而为”的服务,用英文表示是“best effort delivery”,意思是“尽最大努力交付”。也有的简单地用“best-effort”表述。不管用什么样的语言表述,它只想说明IP协议的一个重要的设计思想,即分组在使用IP协议通过多个路由器逐跳传输时,除非在IP分组出现传输差错、拥塞或路由器队列溢出等情况下,会出现IP分组丢失的现象,在正常情况下IP协议将尽最大能力,通过路由算法选择不同的传输路径来正确地传输IP分组。

(2) 以下两点是IP协议不能保证做到的。

① 不能够保证从源主机发送的属于一个报文的多个IP分组,在传送到目的主机时不出现乱序、重复与丢失的现象。

② 不能够保证从源主机发送的IP分组在规定的时间内传送到目的主机。

最初的IP协议文本只定义了IP分组、分组头与标准分类IP地址结构,没有考虑如果出现分组丢失的处理措施,设计者认为保证高层数据传输的准确性与完整性应该由传输层协议来解决。在这之后尽管也增加了ICMP,但是ICMP只能对出现的一些分组丢失的理由通报给发送主机,而本身不采取任何补救措施。IP协议的这个特点看起来太简单、不可靠,但是从几十年应用的结果看,正是这种“用简单方法处理复杂问题”的思路,使得IP协议能够适应复杂的应用需求,在推动Internet发展上起到了非常重要的作用。

问题5-3:最初的IPv4协议主要包括哪些内容?

如果从网上检索出RFC791以及相关的文档可以知道,1981年发布的IPv4文档RFC791只对IPv4分组的结构,32位标准分类的IPv4地址、寻址,以及分片与重装做出了相关的规定。文档中用数据包描述IP协议数据单元。图5-2给出了RFC791包括的主要内容的示意图。

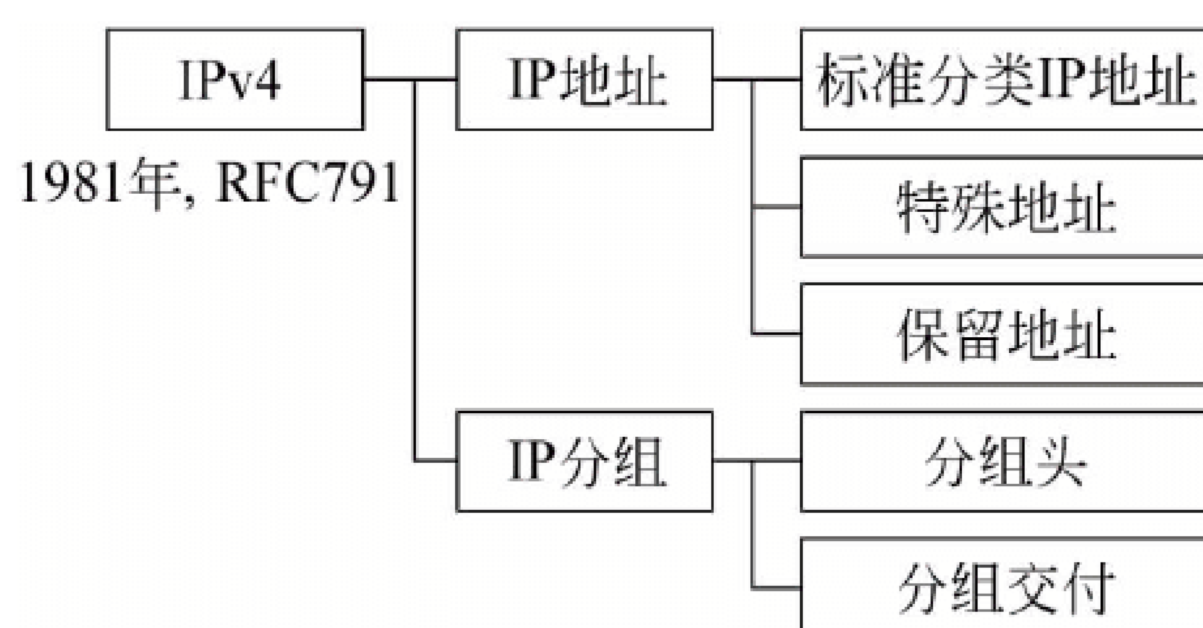


图5-2 RFC791包括的主要内容

整个IP协议的讨论就是在这样一个基础上展开的,但是无论后人做了什么样的补充和完善,IP协议的框架一直没有大的改变。

问题5-4:IPv4协议的缺陷主要表现在哪几个方面?

1. IPv4协议的设计思想

用简单的方法解决复杂问题,用“尽力而为”的服务去应对互联网络中存在的各种复杂的问题,这是IPv4协议的成功之处,也是IPv4掣肘的地方。IPv4在随着计算机网络规模的不断扩大的过程中也逐渐暴露出它的缺陷。

2. IPv4存在的缺点

IPv4存在的问题主要表现在以下几个方面。

(1) 标准分类地址的利用率低,地址数量不能够满足网络规模不断扩展的需要。

(2) 随着网络结构越来越复杂,路由选择算法的研究显得越来越困难。

(3) IPv4 协议对分组传输可靠性没有提供任何保障措施。

(4) IPv4 协议不支持多播传输。

(5) IPv4 协议不能够保证分组传输的服务质量。

(6) IPv4 协议对网络安全问题没有提出对策。

在 20 世纪 70 年代提出的 IPv4 协议,在今天看出问题是非常自然的事,这些年研究人员也正是针对这些问题不断地提出各种完善和补充的意见。

问题 5-5:如何认识 IP 协议发展与演变的过程?

图 5-3 给出了从 1980 年到 2000 年的 20 年时间中 IPv4 协议研究与发展的过程。

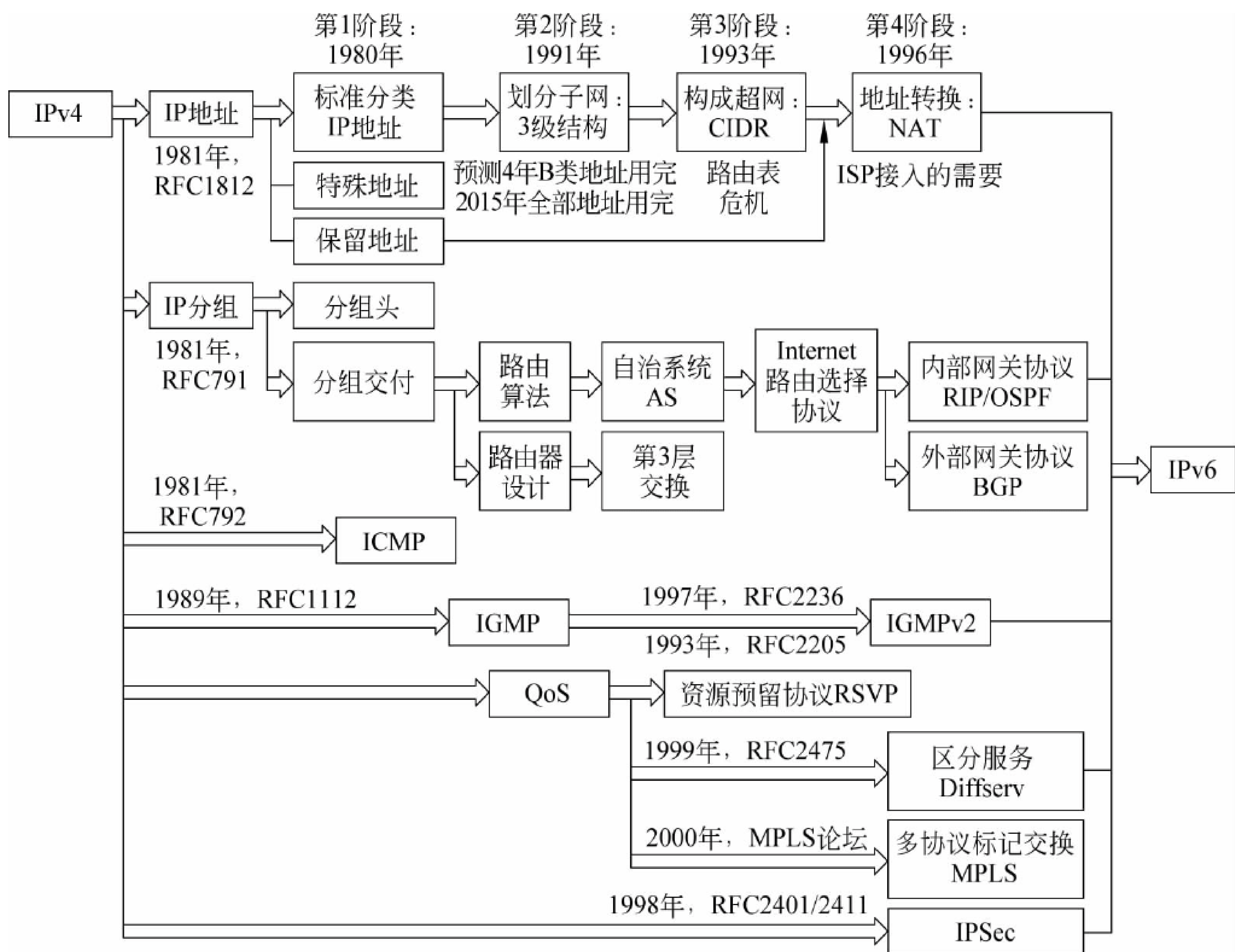


图 5-3 IPv4 协议的研究与发展的过程

(1) 描述 IPv4 协议的最早版本 RFC791 是 1981 年公布的,那个时候 Internet 规模很小,计算机网络主要用于科研与部分参与研究的大学,在这样的背景下产生的 IPv4 协议,不可能适应以后 Internet 规模的扩大和应用范围的扩张,因此修改和完善是必然的。

(2) IPv4 协议发展的过程可以从不变和变化的两个部分去认识。IPv4 协议中对于分组结构与分组头结构的基本定义是不变的;变化的部分可以从 IP 地址处理方法、分组交付需要的路由算法与路由协议,以及为提高协议的可靠性、服务能力与安全角度增加的补充协议等三个部分来认识。

(3) 凡事都有一个限度。当网络规模继续扩大,应用继续深入,这些补充协议已经不能够从根本上解决问题时,就需要彻底考虑重新设计新的协议,这就导致了 IPv6 协议的研究与应用。这是一个很自然的技术发展轨迹。网络技术人员需要了解和认识这个规律。

问题5-6：讲授网络层需要注意哪些问题？

1. 重要性

随着Internet技术的发展,基于IP协议的网络应用成为网络技术研究与软件开发的一个重要基础。人们对未来网络技术的应用前景的描述是:"Everything over IP,IP over Everything"。因此认真学习网络层的基本概念,掌握IP协议的基本内容,对于学生掌握TCP/IP的主要内容,对于网络课程的学习都是十分重要的。本章的内容从整体安排上分量较重,篇幅较多,学习的难度较大。但是无论是网络工程师考试、研究生入学考试,还是就业应聘考核,这一章都占有较重的分量。

2. 复杂性

在网络规模很小的背景之下研究和制定的IP协议,这些年以来研究人员在不断地补充和完善,这就导致IP协议的内容多、变化快、内部结构复杂。初学者在学习本章内容时容易理不清头绪,在学习时感到内容凌乱,不好掌握。所以教师在开始本章教学时,需要对整个章节的内容的结构与教学的安排做一说明,以减少学生学习的盲目性,提高学习质量。

3. 解决办法

作者总结的图5-3也正是希望从一开始就为初学者学习IP协议建立一个整体的概念,然后一步一步地深入学习和掌握。鉴于从知识与技能的掌握,以及各种考核的需要,作者在习题解析与习题中安排了较多的内容,以适应学习的需要。

问题5-7：IP协议的特点是什么？

认识IP的特点对了解Internet基本工作原理十分重要。IP协议的特点主要表现在以下几个方面。

(1) IP协议是一种无连接、不可靠的分组传送服务的协议,提供的是一种"尽力而为"的服务。

① 无连接意味着IP协议并不维护IP分组发送后的任何状态信息。每个分组的传输过程是相互独立的。

② 不可靠意味着IP协议不能保证每个IP分组都能够正确、不丢失、不重复和顺序到达目的结点,如果出现校验和错误就立即丢弃该分组,不提供端-端可靠性保证和逐跳确认,以及流量控制与排序等服务。

分组通过Internet的传输过程是十分复杂的,IP协议设计的重点应该放在系统的适应性、可扩展性与可操作性上,而在分组交付的可靠性方面只能做出一定的牺牲。

(2) IP协议是点-点的网络层通信协议。

网络层需要在Internet中为通信的两个主机之间寻找一条路径,而这条路径通常是由多个路由器、点-点链路组成。IP协议要保证数据分组从一个路由器到另一个路由器,通过多跳路径从源结点到达目的结点。因此,IP协议是针对源主机-路由器、路由器-路由器、路由器-目的主机之间的数据传输的点-点线路的网络层通信协议。

(3) IP协议屏蔽了互联的网络在数据链路层、物理层协议与实现技术上的差异。

作为一个面向Internet的网络层协议,它必然要面对各种异构的网络和协议。在IP协议设计中,设计者就充分考虑了这一点。互联的网络可能是广域网,也可能是城域网或局域网。即使都是局域网,它们的物理层、数据链路层协议也可能不同。协议的设计者希望使用

IP 分组来统一封装不同的网络帧。通过 IP 协议，网络层向传输层提供的是统一的 IP 分组，传输层不需要考虑互联网络在数据链路层、物理层协议与实现技术上的差异，IP 协议使得异构网络的互联变得容易了。图 5-4 体现出 IP 协议对低层协议与实现技术差异的屏蔽作用。

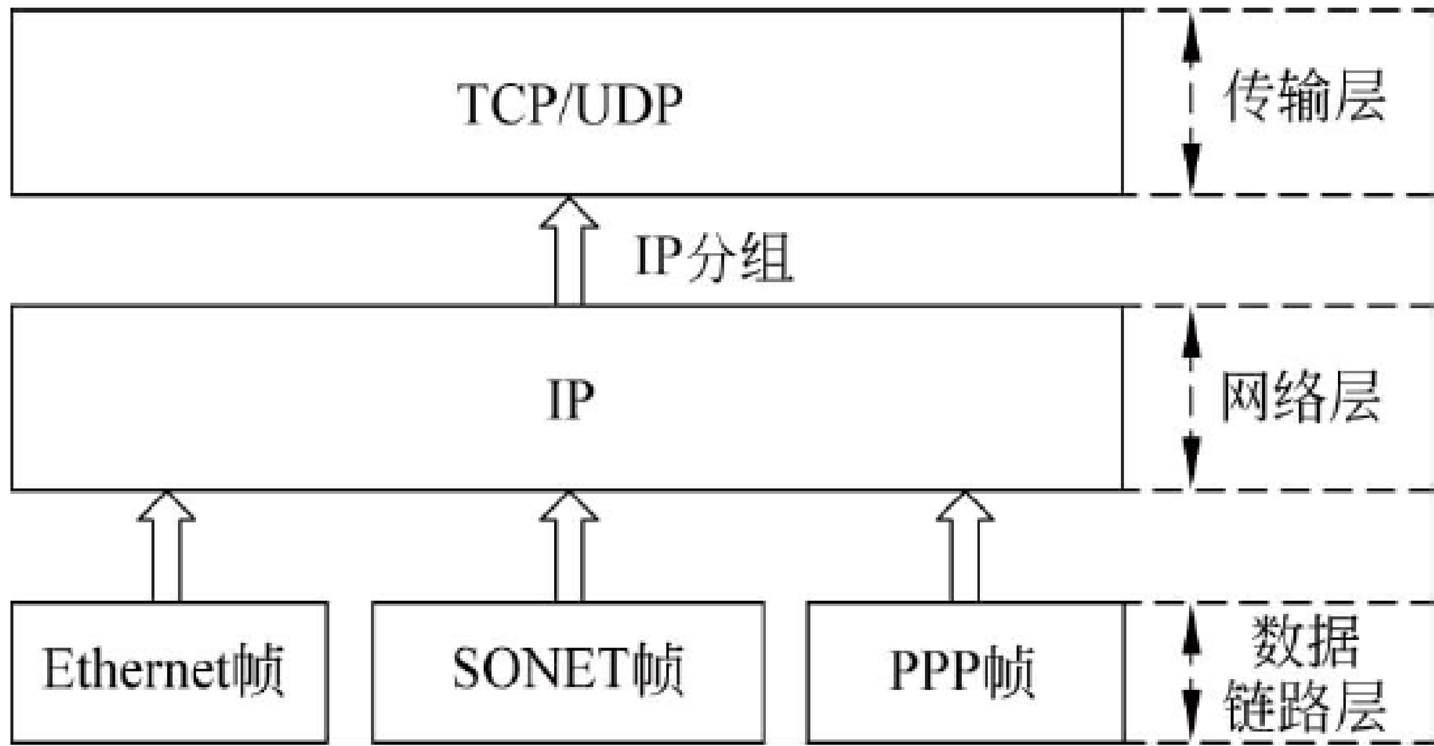


图 5-4　IP 协议对低层协议与实现技术差异的屏蔽作用

问题 5-8：如何认识 IPv4 与 IPv6 报头结构的特点？

在 IP 协议的讨论中，"分组"也称为"报文"，"分组头"也称为"报头"。早期的 RFC 文档使用的术语是"报文"与"报头"。图 5-5(a)与图 5-5(b)分别给出了 IPv4 与 IPv6 报头的结构示意图，从这两张图中可以直观地看出二者的区别。

IPv4 与 IPv6 报头结构的区别主要表现在以下几点。

(1) 组成 IPv4 报头的字段(包括选项)共有 14 个，而组成 IPv6 基本报头的字段数量已经减少到 8 个。IPv6 报头取消了 IPv4 报头的字段是：报头长度、标识、标志、片偏移、头校验和、选项与填充域等(如图 5-5(a)中阴影部分所示)；IPv6 报头增加的字段是流标识(如图 5-5(b)中阴影部分所示)。

(2) IPv4 报头长度是可变的，IPv6 报头长度是固定的。因此，IPv6 报头可以取消 IPv4 报头的"报头长度"字段。

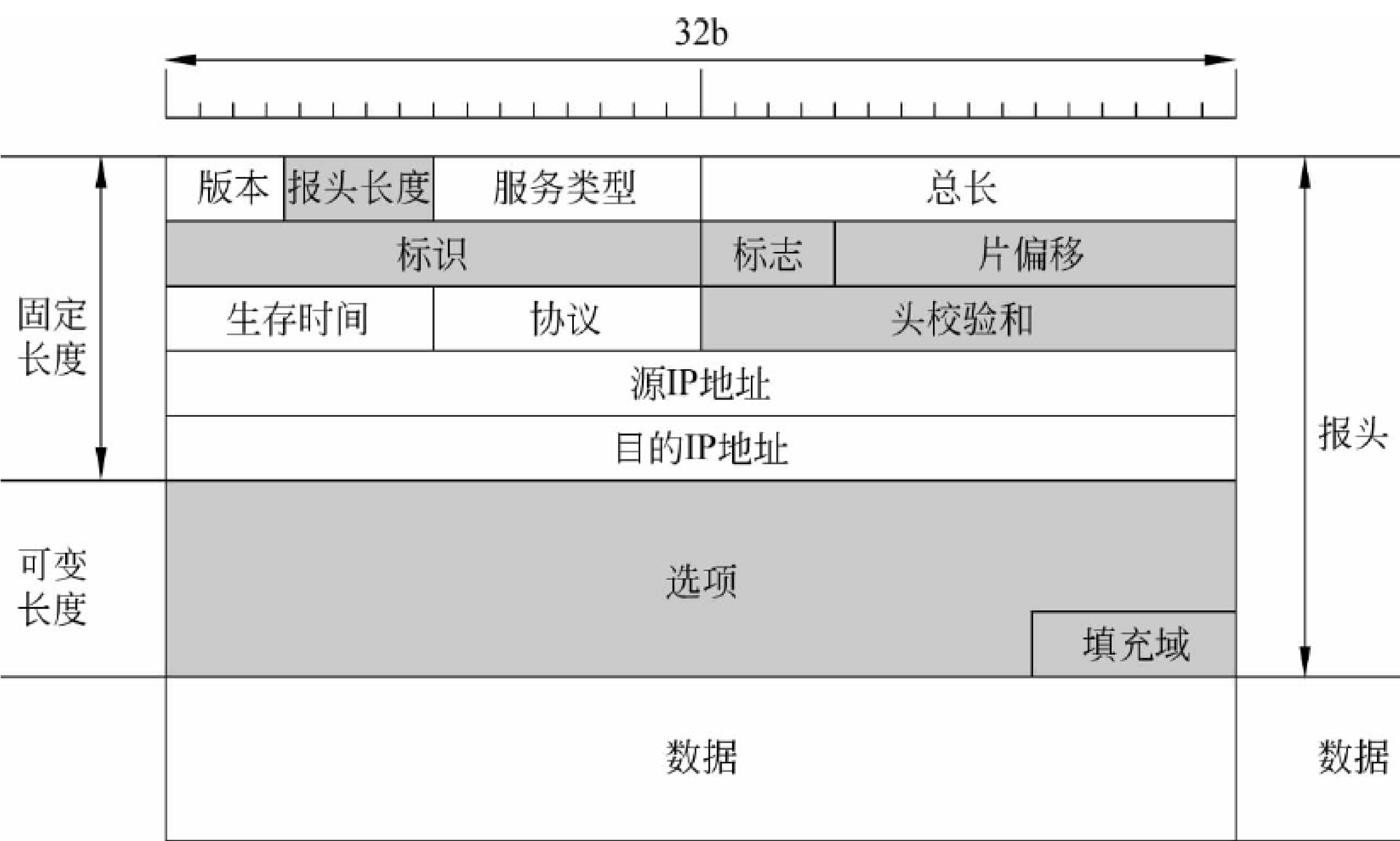
(3) 由于目前大量使用 Ethernet 的 MAC 层与 PPP，在帧处理过程中低层协议中都采用了数据传输差错校验与差错控制机制，因此取消 IPv4 协议中"头校验和"不会对数据传输可靠性产生很大的影响，减轻了路由器的工作负荷，缩短了路由器的转发延时，提高了传输网的工作效率。

(4) IPv6 用"载荷长度"字段取代了 IPv4 的"总长度"字段。IPv4 的"总长度"字段值指包括报头在内的报文总长度，而 IPv6"载荷长度"字段的数值只表示报文的有效载荷长度。

(5) IPv6 用"通信流类型"字段取代了 IPv4"服务类型"字段；用"跳数限制"字段取代了"生存时间"字段；用"下一个报头"字段取代了"协议"字段。IPv6 用"扩展报头"字段取代了"选项" 字段。

(6) IPv6 协议规定源主机可以通过"路径 MTU 发现"报文，了解转发路径中 MTU 的值，可以不发送大于 MTU 长度的分组，转发路由器不执行分片操作。因此，IPv6 报头不需要有 IPv4 报头中的支持拆分的"标识""标志"与"片偏移"字段。

表 5-1 给出了 IPv6 与 IPv4 报头对应字段作用和功能的比较。



(a) IPv4



(b) IPv6

图5-5　IPv4与IPv6报头结构的比较

表5-1　IPv6与IPv4报头对应字段作用的比较

IPv4报头字段	作　　用	IPv6报头字段	作　　用
版本 (version) 4b	协议版本号 (数值为4)	版本 (version) 4b	协议版本号 (数值为6)
报头长度 (header length) 4b	以4字节为单位的报头长度数,包括报头选项		
服务类型 (type of service) 8b	指定优先级、可靠性与延迟参数	通信流类型 (traffic class) 8b	允许源主机高层用户根据应用的需求,确定IP分组的不同的处理类型或不同优先级;默认值为0

续表

IPv4 报头字段	作　　用	IPv6 报头字段	作　　用
		流标记 (flow label) 20b	路由器根据流标记值,在转发时采用不同策略;默认值为 0
总长度 (total length) 16b	以字节为单位的报文总长度	载荷长度 (payload length) 16b	表示包括报文的有效载荷长度
标识 (identification) 16b	标识属于同一个报文的不同分片		
标志(flag) 3b	表示报文"还有分片"与"不能分片"		
片偏移 (flagment offest) 13b	分片的偏移量		
		下一个报头 (next header) 8b	如果报文有附加的扩展报头,该字段之后为下一个扩展报头;如果没有,则用于标识传输层的协议类型是 TCP 或 UDP
生存时间 (time to live) 8b	报文在网络中以秒为单位的寿命	跳数限制 (hop limit) 8b	报文在网络中经过路由器最多转发的次数
协议(protocol) 8b	高层协议的类型		
头校验和 (header checksum) 16b	只校验报文头,不包括数据部分		
源地址 (source address) 32b	发送方的 IPv4 地址	源地址 (source address) 128b	发送方的 IPv6 地址
目的地址 (destination address) 32b	接收方的 IPv4 地址	目的地址 (destination address) 128b	接收方的 IPv6 地址
选项 (options) 24b	用户可以选择的,用于控制与测试目的报头选项		
填充位 (padding)	用于保证报头是 32b 的整数倍		

问题 5-9:如何认识 IPv6 基本报头的特点?

深入理解 IPv6 报头的基本内容与特点,需要与 IPv4 做一个对比和分析。

1. 版本(version)字段

"版本"字段的意义与 IPv4 相同,版本字段值为 6,表示使用的是 IPv6 协议。需要注意

的是,在 Ethernet 帧封装时,帧头的 EtherType 字段在网络层使用 IPv4 协议时,值为 0x800;在使用 IPv6 协议时,值为 0x86DD。

2. 通信流类型(traffic class)字段

通过"通信流类型"字段,源主机可以通过设定不同的数值,来区分不同分组的类型或优先级。例如,通信流类型字段值为 0~7 时,表示在拥塞发生时允许延时处理;字段值为 8~15 时,表示是优先级较高的实时业务,需要用固定速率传输。传输路径中的转发路由器可以根据通信流类型字段值,来提供不同类型的服务。源主机不选择区分分组优先级与不同类型传输服务时,通信流类型字段的默认值为 0。

3. 流标记(flow label)字段

"流标记"字段长度为 20b。这里所说的"数据流"是指从某个特定的源主机向另一个特定的目的主机发送的单播或组播数据分组序列。当源主机需要转发路由器对该报文序列采用特殊的服务时,需要使用流标记字段。源节点不使用流标记字段功能,要将流标记字段值置 0。

IPv6 协议采用流标记字段可以实现资源预留协议(RSVP)与实时流协议/实时传输协议(RTP/RTCP)。而 RSVP 与 RTP/RTCP 提供的服务对于要求固定带宽、固定延迟、实时性要求很高的音频、视频等应用至关重要。源主机根据实际应用需求对属于同一个数据流的数据分组,设定一个流标记字段的值。属于同一个数据流的分组具有相同的源与目的 IPv6 地址、相同的流标记字段值。路由器可以通过简单地区分具有相同流标记字段值的数据分组,来提供保证 QoS 的转发服务。

RFC2460 对通信流类型、流标记字段的使用没有做出明确的定义。目前,IPv6 的流标记字段的相关问题正在实验改进之中。使用通信流类型、流标记字段,需要 IPv6 路由器与主机通过 RSVP 或 RTP/RTCP 的协助,才能够实现 QoS 服务功能。

4. 载荷长度(payload length)字段

由于 IPv6 报头长度固定,因此不需要像 IPv4 那样专门设置一个"报头长度"字段去说明。"载荷长度"字段值用来表示报文的有效载荷长度。

IPv6 载荷长度字段长度为 16b,可以表示长度最大为 65 535B 的有效载荷。如果有效载荷长度超过最大长度,则载荷长度字节置 0,需采用扩展报头中逐跳(hop-by-hop)选项中的"超大有效载荷(Jumbo Payload)"来处理。

5. 下一个报头(next header)字段

"下一个报头"字段长度为 8b,它的作用与 IPv4 报头中的"协议(protocol)"相似,并且 IPv6 的"下一个报头"字段的值与 IPv4 报头中的"协议(protocol)"字段的值都是由 RFC1700 定义的。不同的是:IPv4 的"协议"字段的值表示的是高层协议的类型,而 IPv6 "下一个报头"字段的值表示紧跟在 IPv6 基本报头之后的字段是扩展报头,还是 TCP 或 UDP 的数据。表 5-2 给出了常用的 6 种"下一个报头"字段值、对应的扩展报头类型与扩展报头的基本功能。

同时,需要注意以下几点。

(1) 如果"下一个报头"字段值为 6,表示紧跟在 IPv6 基本报头之后的字段是 TCP 数据。

表 5-2 “下一个报头”字段值、扩展报头类型与功能

扩展报头类型	下一个报头值	功 能 讨 论
逐跳选项报头（Hop-By-Hop option header，HBH）	0	主要用于网管或网络软件调试时要求路由器做特殊处理的信息（如路由器不能识别选项类型时应采取的处理方法、路由器警告或超大有效载荷长度报文），也是唯一的一个需要传输路径上所有路由器都要处理的扩展报头
目的地选项报头（Destination Option Header，DOH）	60	携带的信息仅需要目的主机处理
路由探头（Routing Header，RH）	43	指定源路由，用于网络故障诊断测试与移动 IPv6 解决迂回路由
分片报头（Fragment Header，FH）	44	表示源主机已经对报文进行分片，由目的主机将分片重新组装成报文，不要求转发路由器处理
认证报头（Authentication Header，AH）	51	保证报头的安全性，由源主机与目的主机，或源安全网关与目的安全网关处理
封装安全载荷报头（SIPP Encap Security Payload Header，ESP）	50	保证有效载荷的安全性，由源主机与目的主机，或源安全网关与目的安全网关处理

（2）如果“下一个报头”字段值为 17，表示紧跟在 IPv6 基本报头之后的字段是 UDP 数据。

（3）如果“下一个报头”字段值为 59，表示该报头是最后一个报头，并且报文也没有携带 TCP/UDP 数据。

6．跳数限制（hop limit）字段

跳数限制字段长度为 8b，表示 IPv6 分组可以通过的最多的转发路由器的数量。IPv6 的跳步限制字段与 IPv4 的 TTL 字段非常相似。分组每经过一个路由器时，数值减 1。当跳数限制字段的值减为 0 时，路由器丢弃该报文，并向源节点发送 ICMPv6“超时”报文。

在介绍了 IPv6 基本报头结构之后，可以给出一个例子来对这个问题做一个小结。下面是用 Network Monitor 截获的一个实际的 ICMPv6 协议回送请求报文的基本报头。

```
Frame:Base frame properties
    …
Internet Protocol:
    version= 6(IPv6)
    traffic class= 0x0(0)
    flow label= 0x0(0)
    playload length= 0x0028(40)
    next header= 0x003A(58,ICMPv6)
    hop limit= 128
    source address= 2001::160:16ff:cd02:3d02
    destination address= 2001::32a:16ff:2:35
    …
```

从这个基本报头可以看出:

(1) traffic class = 0x0、flow label = 0x0,表示报文使用默认的通信类型标识与流标记。

(2) payload length = 0x0028,换算成十进制是40,表示基本报头长度是标准的40B。

(3) next header = 0x003A,换算成十进制是58,表示扩展报头是ICMPv6协议报头。

(4) hop limit = 128,表示报文最多经过128个中间路由器转发。

(5) source address= 2001::160:16ff:cd02:3d02,表示的是源IPv6地址。

(6) destination address = 2001::32a:16ff:2:35,表示的是目的IPv6地址。

问题5-10:如何认识IPv6扩展报头的特点?

认识IPv6扩展报头的特点,需要与IPv4做一个比较。

1. 为什么要设置扩展报头

在IPv4网络中,IP报文的报头在经过每个中间转发路由器时,路由器都必须检查报头的"选项"是否存在。如果存在就必须对长度可变的分组头进行处理。这种做法势必要增加路由器处理报头的计算负荷,降低路由器转发IPv4报文的效率。在IPv6中,报头只保留路由器必须处理的内容,并且长度固定,而将"选项"放到了扩展报头中。扩展报头由源主机按需要添加。每个转发IPv6报文的中间路由器,只处理固定长度的基本报头,而唯一需要处理的扩展报头就是"逐跳选项报头"。因此,这种做法必然会提高路由器处理IPv6报头的速度,缩短路由器转发IPv6报文的延迟时间。

了解IPv6扩展报头的特点,需要注意以下几个问题。

(1) 如果有多个扩展报头,则按从"逐跳选项报头""目的地选项报头""路由报头""分片报头""认证报头"到"封装安全载荷报头"的顺序排列。

(2) 每一个扩展报头都是由若干个字节组成,长度各不相同,但是必须是8字节的整数倍,并且扩展报头的第一个字节都应该是"下一个报头"字段;第二个字节是"报头长度",用于标识扩展报头的长度(不包括扩展报头的第一个字节)。

(3) RFC2460并没有对涉及安全的"认证报头"和"封装安全载荷报头"给出定义。"认证报头"和"封装安全载荷报头"是由IETF制定的IPSec协议的"认证头(AH)协议"与"封装安全载荷(ESP)协议"定义。"认证报头"与"封装安全载荷报头"由源主机与目的主机,或者是由源安全网关与目的安全网关来处理。

(4) 一般的IPv6报文并不需要这么多的扩展报头,只是在需要转发路由器或目的主机需要配合做一些特殊处理时,例如在网管软件、网络软件测试与网络故障诊断中,源主机才会添加一个或几个必要的扩展报头。

(5) 获得"下一个报头"字段意义的最新列表的地址是:

http://www.ian6.org/assignments/protocol-numbers。

2. 由"下一个报头"字段组成的指针链

IPv6分组中可以没有扩展报头,可以只有一个扩展报头,也可以有多个扩展报头。

第一种情况:没有扩展报头。

没有扩展报头的情况如图5-6所示。如果IPv6的基本报头中"下一个报头"字段值为6,表示该分组没有扩展报头,它的有效载荷是TCP包的数据。如果"下一个报头"字段值为17,则表示该分组没有扩展报头,它的有效载荷是UDP包的数据。

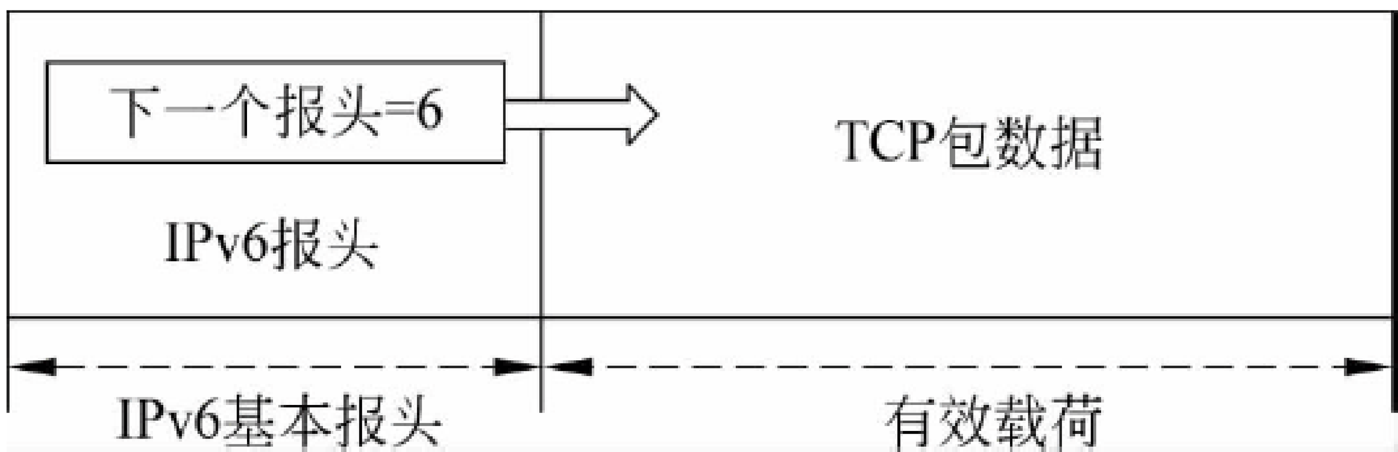


图 5-6　有效载荷是 TCP 数据

第二种情况：有一个扩展报头。

图 5-7 给出了只有一个"路由报头"的 IPv6 报头结构示意图。如果分组只有一个"路由报头"，"下一个报头"字段值为 43，表示该分组的有效载荷第一部分为"路由报头"。在"路由报头"的第一个字节的"下一个报头"字段中，应该指出该分组的有效载荷的数据部分应该交给 TCP 处理，还是交给 UDP 处理。如果"路由报头"的第一个字节的"下一个报头"字段中值为 6，则表示高层协议是 TCP；如果值为 17，则表示高层协议是 UDP。

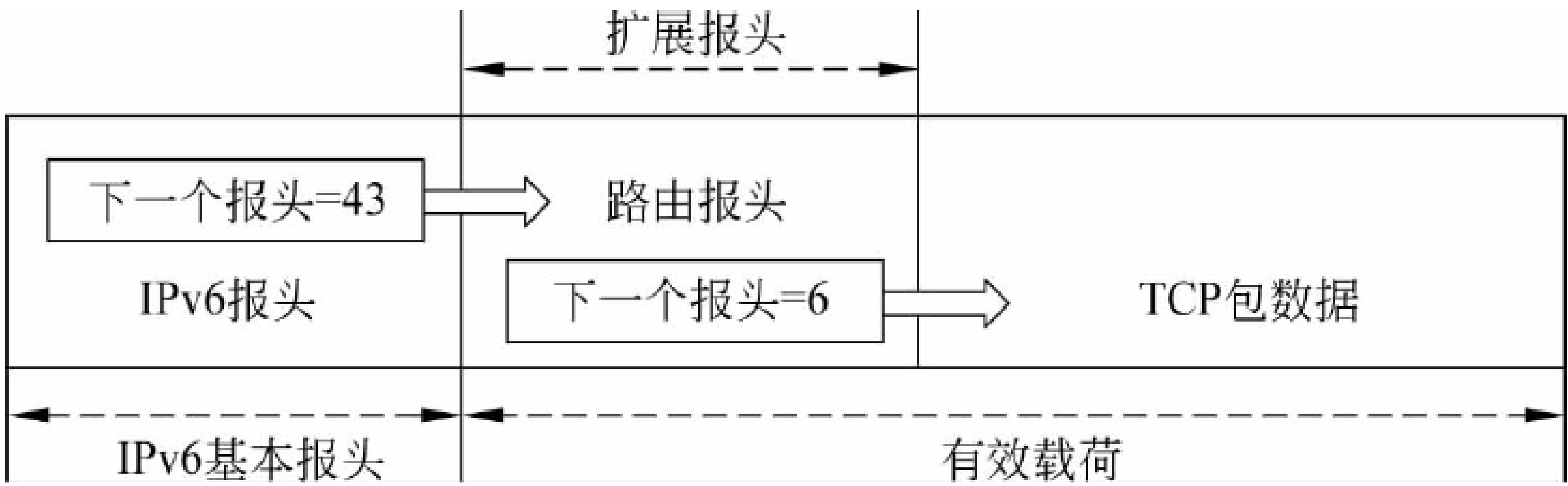


图 5-7　只有一个扩展报头的 IPv6 报头结构

第三种情况：有两个扩展报头。

图 5-8 给出了有两个扩展报头的 IPv6 报头结构示意图。如果分组有两个扩展报头，第一个是"路由报头"，第二个是"认证报头"。那么，IPv6 的基本报头中"下一个报头"字段值仍然为 43，表示第一个扩展报头是"路由报头"。在路由报头之后的第二个扩展报头是"认证报头"，那么"路由报头"的第一个字节的"下一个报头"值应该为 51。"认证报头"之后再没有扩展报头了，那么"认证报头"的第一个字节"下一个报头"值应该为 6 或 17。如果为 6，则表示该分组的有效载荷数据为 TCP 数据。

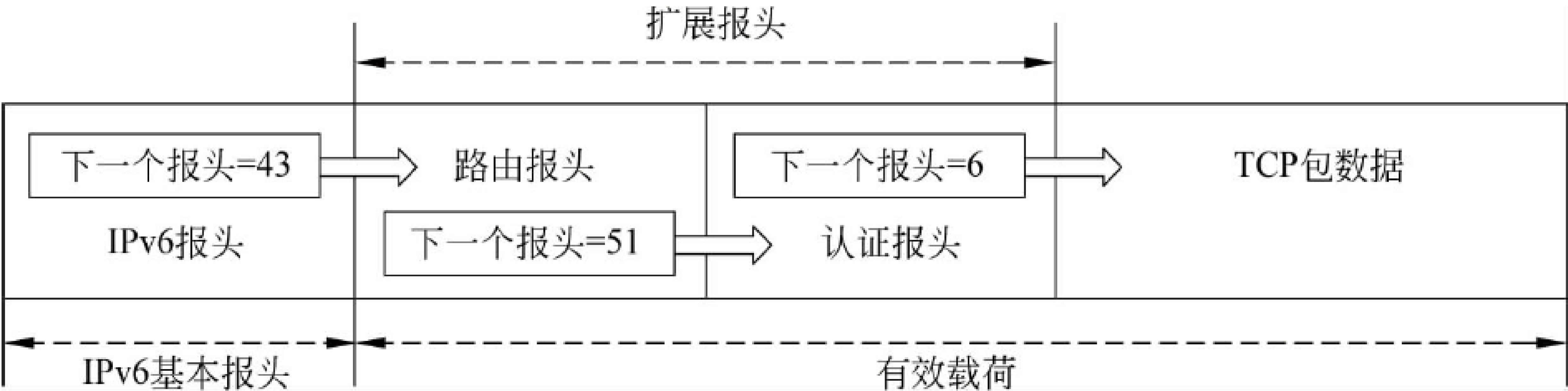


图 5-8　有两个扩展报头的 IPv6 报头结构

从以上例子的分析中可以看出：IPv6 的基本报头中的"下一个报头"字段与扩展报头的"下一个报头"组成了有关扩展报头的指针链表。每个指针表示出紧接着它的下一个扩展报头的类型，最后一个扩展报头的"下一个报头"指出高层协议的类型，其结构如图 5-9 所示。其中，基本报头长度固定为 40B，而扩展报头长度是可变的，扩展字节中的报头长度值标明扩展报头的长度。

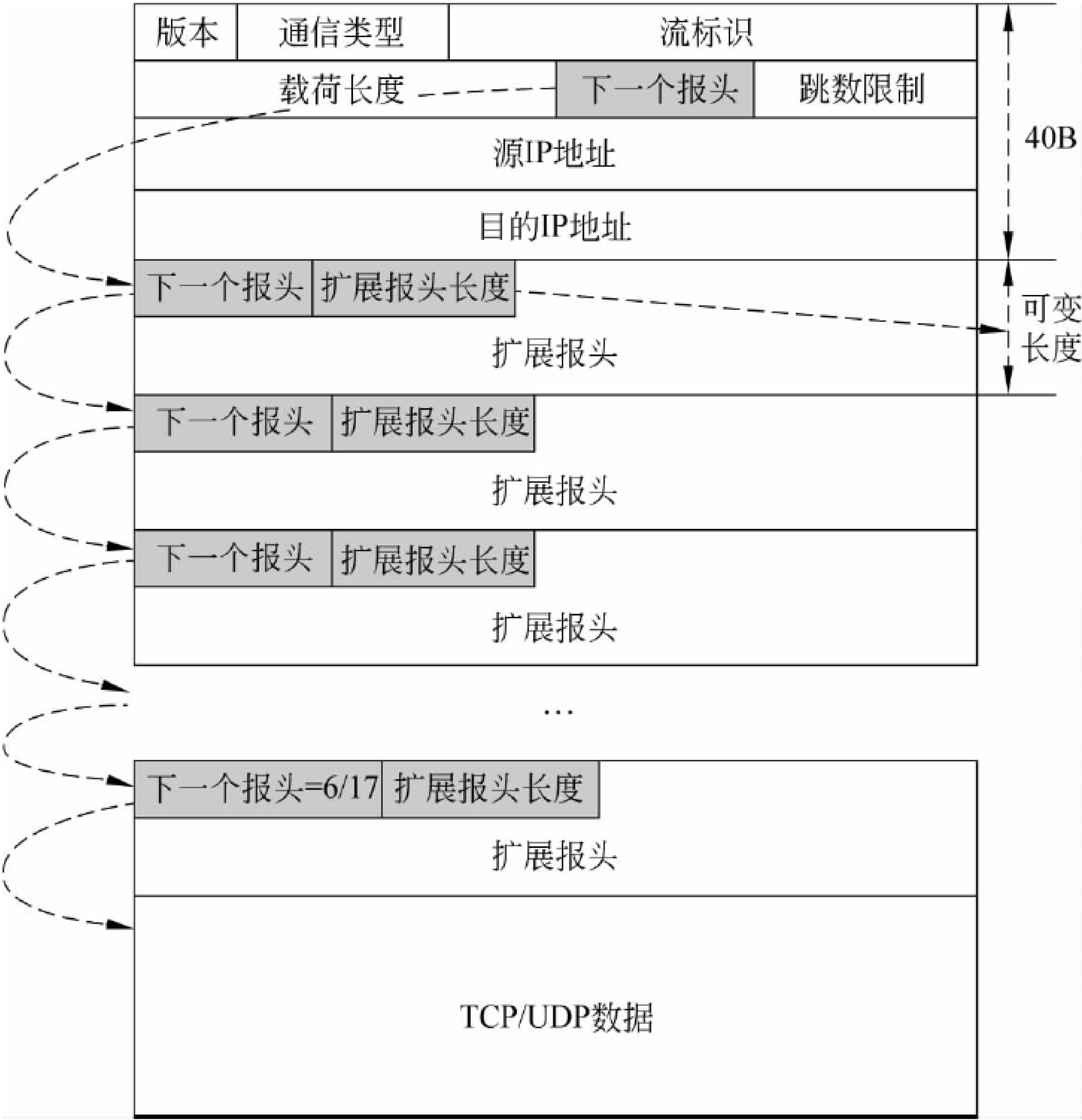


图 5-9 由“下一个报头”字段组成的指针链示意图

3. IPv6 扩展报头选项字段的意义

目前常用的 IPv6 扩展报头主要有 6 个：逐跳选项报头、目的地选项报头、路由报头、分片报头、认证报头与封装安全载荷报头。

(1)“逐跳选项报头”是转发路由器唯一需要处理的一个扩展报头。它携带着网管或网络软件调试要求路由器特殊处理的信息。

(2)“目的地选项报头”选项表示该报文携带着只能被目的主机检查的信息。

(3)“路由报头”选项与 IPv4“源路由”选项作用类似。路由报头包含报文在从源节点到达目的节点的过程中,需要经过的一个或多个中间转发路由器的地址列表。

(4)“分片报头”的用途是表示源主机发送了一个大于 MTU 长度的报文。源主机在发送之前,通过发送“路径 MTU 发现”报文,找出沿该路径到目的主机的所有链路中链路 MTU 最小的一个。如果源主机发现待发送的报文长度大于 MTU,源主机对原始的报文已经进行了分片,并将每个分片作为独立的分组发送出去。这些分片到达目的主机之后,由目的主机再重新组装起来,转发的 IPv6 路由器不需要进行分片处理。

(5)“认证报头”选项提供对需要保护的数据,进行数据认证、数据完整性检查与反重放攻击的保护。认证报头选项由源与目的主机,或源与目的安全网关处理,以提高端-端通信的安全性。

(6)“封装安全载荷报头”选项 ESP 可以与认证报头 AH 结合起来使用,也可以单独使用。封装安全载荷报头是由源与目的主机,或者由源与目的安全网关处理。

4. IPv6扩展报头选项的使用举例

我们以“逐跳选项报头”的使用为例，来进一步说明IPv6扩展报头选项的作用。理解“逐跳选项报头”的作用，需要注意以下几个问题。

1）“逐跳选项报头”结构

逐跳选项报头结构如图5-10所示。逐跳选项报头由8b的下一个报头、8b的扩展报头长度与选项三部分组成。选项长度是8b的整数倍。扩展报头长度表示的是选项长度以8b为单位的数值(图5-10中所示的N值)。



图5-10　逐跳选项报头结构

2）选项类型的意义

(1) 8b的选项类型最高两位值表示当前处理选项的路由器不能识别“选项类型”时，应该采取的处理方法。

00——可以跳过这个选项。

01——丢弃该分组。

10——丢弃该分组；如果目的IPv6的地址为单播或多播地址，向源节点发送一个ICMPv6“参数错误”报文。

11——丢弃该分组；如果目的IPv6的地址不是多播地址时，向源节点发送一个ICMPv6“参数错误”报文。

(2) 8b的选项类型最高的第三位值表示当前处理选项的节点是否能够改变分组到达目的节点的路由。

第三位值为0时，表示当前处理选项的节点不能改变路由。

第三位值为1时，表示当前处理选项的节点可以改变路由。

(3) 超大有效载荷选项。

在“逐跳选项报头”使用中有两种情况需要注意，那就是超大有效载荷与路由器警告。

超大有效载荷的选项类型在RFC2675中做了定义，超大有效载荷的选项类型的结构如图5-11所示。超大有效载荷的选项类型值为192，选项长度值为4。

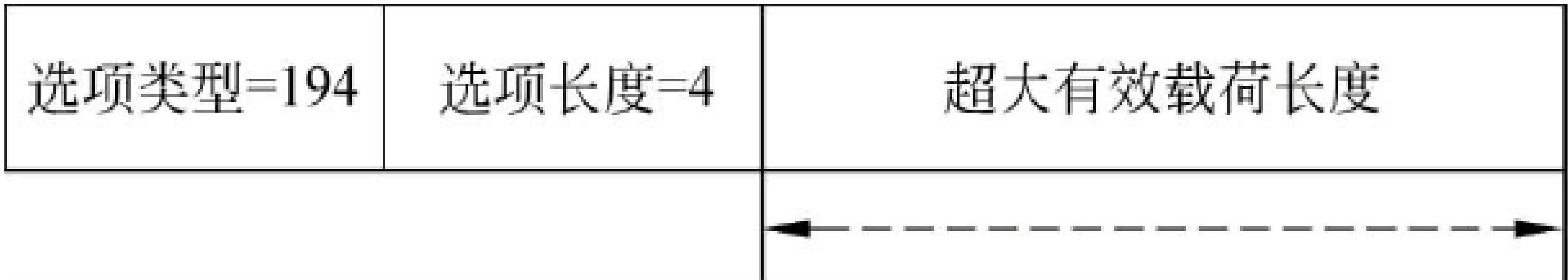


图5-11　超大有效载荷的选项类型的结构

如果使用超大有效载荷的选项，则IPv6报头中的有效载荷长度不再用IPv6的有效载荷长度字段表示，而是用超大有效载荷的“超大有效载荷长度”字段表示，单位是字节。由于超大有效载荷选项长度为4B，因此它可以表示最大的有效载荷长度为4 294 967 295B。当

路径 MTU 超过 65 535B 的有效载荷长度时,就可以选择该选项来传输超大有效载荷的分组。

(4) 路由器警告选项。

路由器警告选项类型在 RFC2711 中做了定义,路由器警告的选项类型的结构如图 5-12 所示。路由器警告的选项类型值为 5,选项长度值为 2,选项数据值为 0。

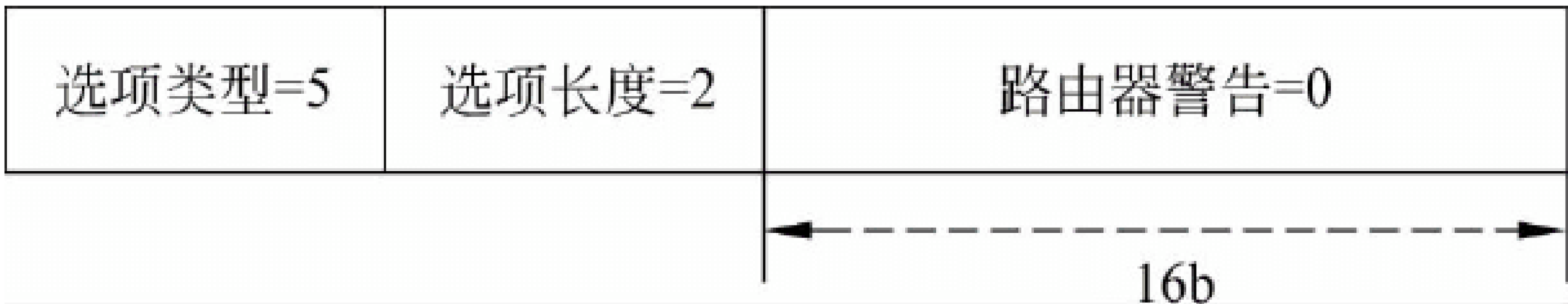


图 5-12 路由器警告的选项类型结构

路由器警告的选项用于多播侦听发现(MLD)和资源保留协议(RSVP)。由于路由器警告的选项类型值为 5,该选项如果不被处理它的节点所识别,那么可以跳过该选项,但选项在传输过程中不允许改变。

下面给出一个通过 Network Monitor 截获的"逐跳选项报头"带有"路由器警告"选项的简化报头结构。

```
Frame:Base frame properties
Ethernet:
    Etype= IPv6
IPv6:
    hop opts,proto= ICMP,len= 24
    version= 6(0x6)
    traffic class= 0(0x0)
    flow label= 0(0x0)
    payload length= 32(0x20)
    next header= 0(hop- by- hop options header)
    hop limit= 1(0x1)
    source address= 2001::160:16ff:cd02:3d02
    destination address= 2001::32a:16ff:2:35
    hop- by- hop options header
        next header= 58(ICMPv6)
        length= 0(0x0)
        type= 5(router alert option)
        0 0 _ _ _ _ _ _= skip option if not recognized
        _ _ 0 _ _ _ _ _= option data does not change enroute
        length= 2(0x2)
        router alert value= 0(0x0)
        padding(2bytes)
        type= 1(padn)
        0 0 _ _ _ _ _ _= skip option if not recognized
        _ _ 0 _ _ _ _ _= option data does not change enroute
        length= 0(0x0)
    payload:number of data bytes remaning= 24(0x18)
```

ICMP:multicast listener report

在截获的报头结构中,IPv6: router alert option(type = 5)。同时,为了使路由器警告的选项达到8B,按照协议的相关规定增加了2B的padding选项。这样,1B的“下一个报头”、1B的“选项长度”、4B的“路由器警告选项”,再加上2B的padding选项,使得路由器警告的选项长度达到8B。

问题5-11: 如何从路由器分组转发过程认识IPv4与IPv6的区别?

我们可以通过分析路由器转发一个IPv4分组与转发一个IPv6分组的过程,直观地比较IPv4与IPv6协议的优缺点。

1. 转发一个IPv4报文的过程

IPv4路由器的每一个接收端口都处在随时准备接收IP分组的状态。当它接收一个IPv4分组之后立即进入处理过程。IPv4分组接收、处理与转发流程如图5-13所示。

路由器对一个IPv4分组从接收到转发的整个过程,大致经过以下几个步骤。

1) 检验“版本”字段

路由器首先要检测分组头的“版本”字段。如果“版本”字段值为4,表示接收到的是IPv4分组,那么接下来可以按照IPv4协议的规定处理。如果“版本”字段值不为4,那么表示接收到的不是IPv4分组,路由器丢弃该分组,并向源节点发送ICMP“参数问题”分组,报告传输出错。路由器重新回到准备接收下一个分组的状态。

2) 检验“分组头长度”字段

路由器检测“分组头长度”字段。如果分组头长度正确,进入下一个流程。如果分组头长度错误,路由器丢弃该分组,并向源节点发送ICMP“参数问题”分组,报告传输出错。

3) 检验“头校验和”字段

IPv4为了及时发现分组头部在传输过程中出现的错误,设置了分组头校验和。路由器要按照IPv4协议规定的计算方法,对接收到的分组头进行校验和的计算。如果发现头校验和错误,路由器丢弃该分组,并向源节点发送ICMP“参数问题”分组,报告传输出错。路由器重新回到准备接收下一个分组的状态。如果头部校验和正确,路由器进入下一个流程。

4) 检测“TTL”字段

最初人们在设计IPv4生存时间TTL时,是希望在分组中保持一个时间戳,用来限制一个分组在互联网络中的最大生存时间,时间单位是秒。但是在实际使用时,生存时间TTL的值被用来表示该分组最多可以经过几个路由器转发。TTL的初始值由源主机设置,每经过一个路由器,TTL值减1,填入TTL字段。

如果路由器检查出TTL值等于0,路由器丢弃分组,并向源节点发送ICMP“超时”分组。TTL值大于1,路由器将TTL字段值减1后,填入TTL字段;转到下一个处理流程。

5) 处理“报头选项”字段

IPv4协议有一个IP选项(options)字段。如果分组头有“选项”字段时,分组头长度在20～60B,是可变的。路由器软件检查分组头是否有“选项”字段。如果有“选项”字段,就要进行相应的处理。

6) 进行路由选择

路由器路由选择软件使用目的IP地址、源IP地址在路由表中进行查询,来确定转发的

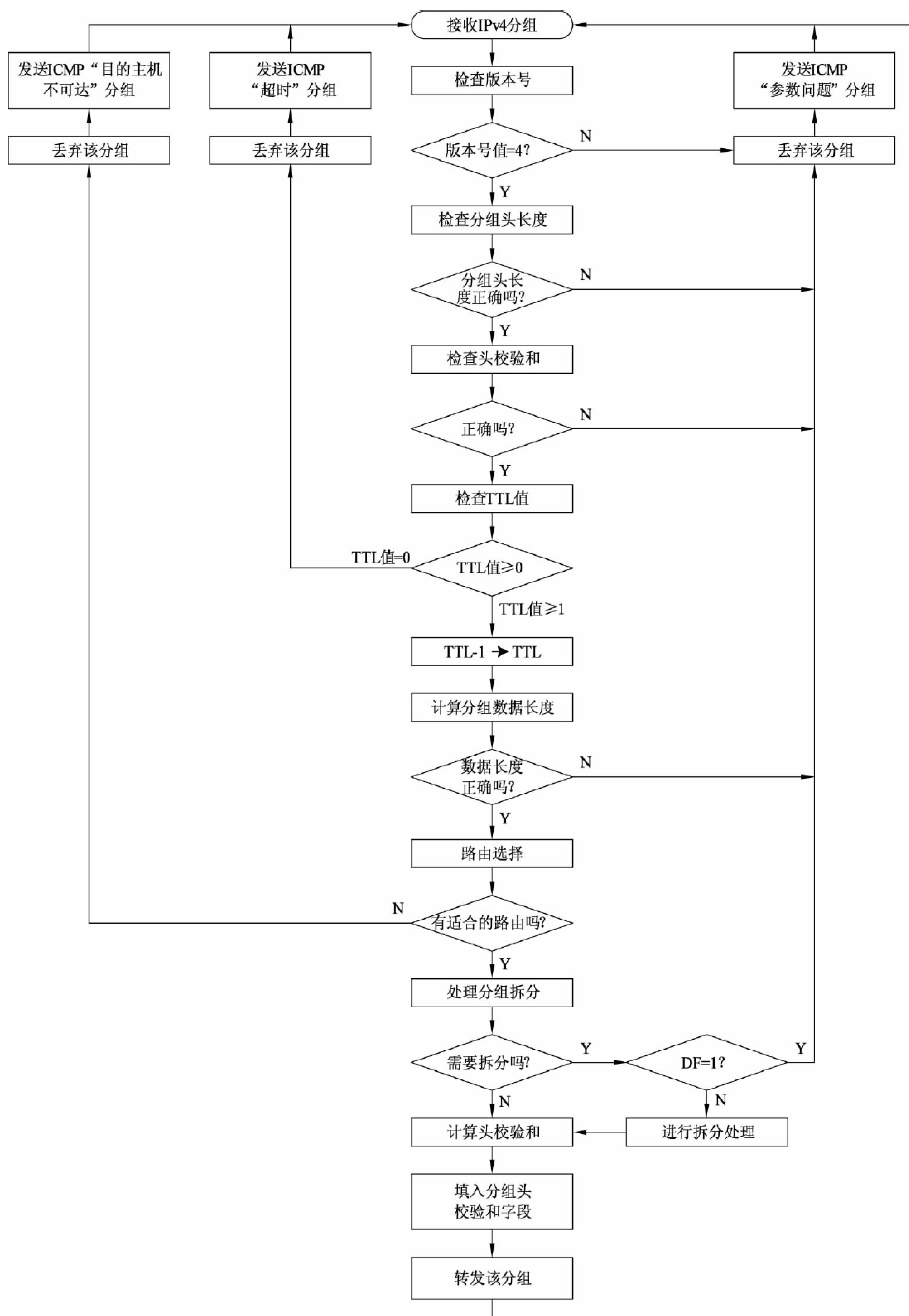


图 5-13 路由器对 IPv4 分组的接收、处理与转发流程

接口，确定下一跳的 IPv4 地址。如果没有找到合适的路由，丢弃该分组，并发送 ICMPv4“目标节点不可达”分组给源节点。

7）处理分组拆分处理

在路由器确定了转发路径时，需要根据下一个通过的网络的最大传输单元的长度，来决定被转发的分组是否要进一步进行拆分。

如果被转发分组的总长度字段的值大于转发接口的最大传输单元 MTU，而分组头中“标志”字段中“不分片(DF)”位的值为 0，表示允许对被转发分组进行拆分，则执行 IPv4 分组的拆分操作。

如果总长度字段的值大于转发接口的 IPv4 最大传输单元 MTU，但是“不分片(DF)”位的值为 1，表示不允许对分组进行拆分。路由器则丢弃分组，并发送一个 ICMPv4 “参数问题”分组给源节点，报告传输出错。

8）计算“头校验和”

在完成分组拆分处理之后，需要重新计算分组新的“头校验和”，并将计算值放置在转发分组的“校验和”字段中。

9）转发分组

路由器在完成以上各步处理之后，可以根据路由选择算法确定的下一跳路由器或目的主机的 IP 地址转发该分组。

需要注意的是，以上给出的是一个 IPv4 路由器在转发一个普通 IPv4 分组时简化的处理过程。这里没有考虑以下几种情况。

(1) IPv4 分组头中 8b 的分组类型(service type)，其中 3b 是优先级(precedence)、4b 是服务类型(Type of Service，TOS)、1b 是保留位。IPv4 分组头没有定义分组类型，定义了 3b 的优先级。在上述流程的描述中，忽略了优先级处理的细节。

(2) IPv4 分组长度是可变的。如果一个有 IPv4 选项(options)字段与填充(padding)字段的 IP 分组，它的分组头要大于 20B。分组头的长度在 20～60B 之间。同时，IPv4 协议规定：IP 分组的分组头长度必须是 4B 的整数倍。如果不是 4B 的整数倍，则由“填充域”通过“添 0”来补齐。上述讨论中忽略了对“选项”字段与“填充”字段处理的细节。

(3) 路由器输入、输出端口都包括网络层、数据链路层与物理层。在确定了下一跳路由器或主机的 IP 地址之后，需要通过 ARP 查询对应于 IP 地址的 MAC 地址。尽管 ARP 属于网络层；ARP 请求与应答分组也是封装在 IP 分组中传输。在路由器软件编程实现的角度也必须有这一部分软件。但是，为了简化原理性描述的过程，在上述流程的讨论中忽略了 ARP 执行的细节。路由器对路由处理的细节将在路由器一节中讨论。

2. 转发一个 IPv6 数据包的过程

路由器接收一个 IPv6 报文之后立即进入处理过程。IPv6 报文接收、处理与转发流程如图 5-14 所示。

下面用同样的方法来分析 IPv6 路由器转发一个 IPv6 报文的过程。

1）检验“版本”字段

路由器首先要检测分组头的“版本”字段。如果“版本”字段值为 6，表示接收到的是 IPv6 报文，那么接下来可以按照 IPv6 协议的规定处理。如果“版本”字段值不为 6，那么表示接收到的不是 IPv6 报文，路由器丢弃该报文，向源节点发送 ICMPv6“参数问题”报文，报

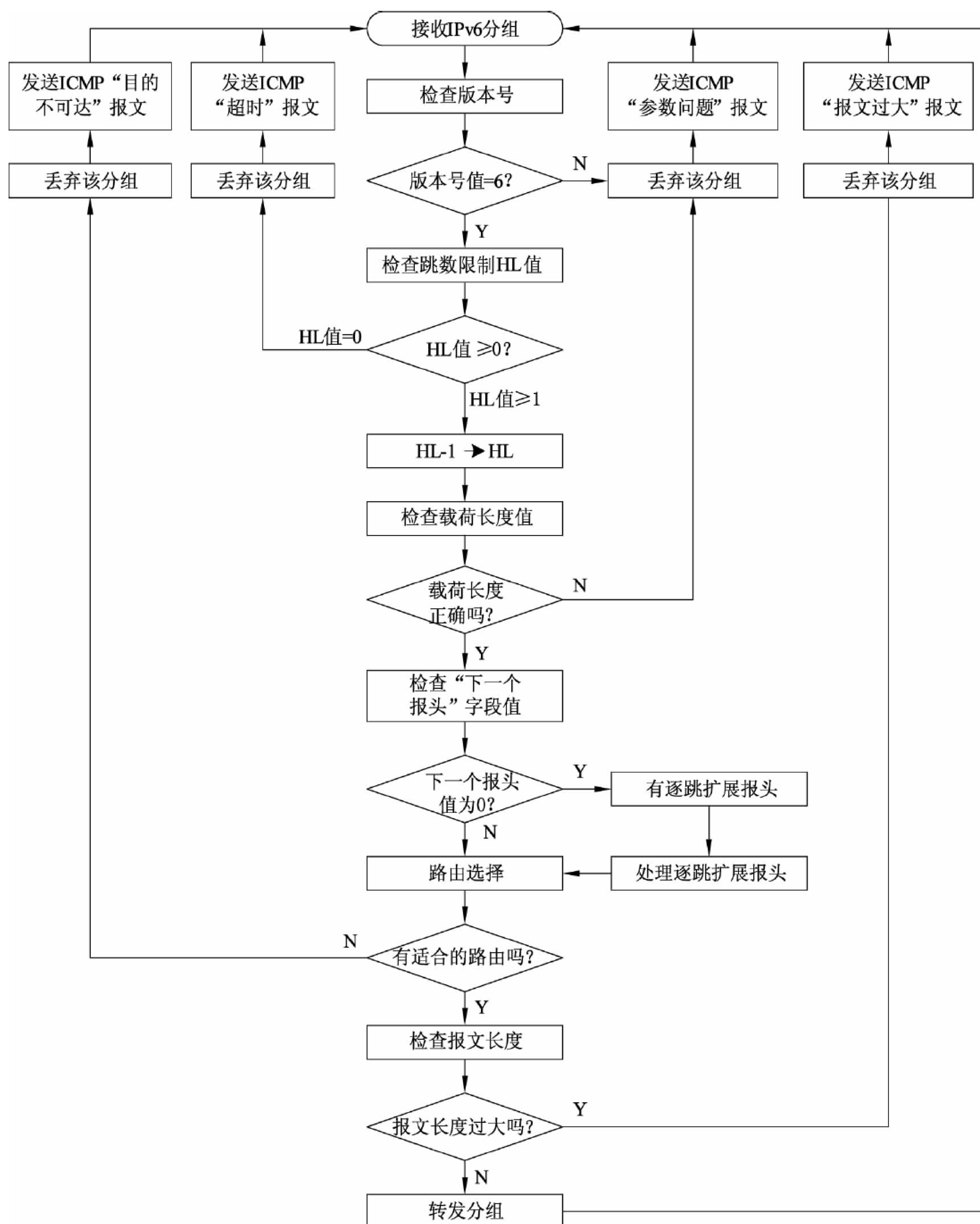


图 5-14　IPv6 报文接收、处理与转发流程

告传输出错。路由器重新回到准备接收下一个分组的状态。

2）处理“跳数限制”字段

跳数限制(Hop Limit,HL)字段的作用与 IPv4 的 TTL 字段相同。报文每经过一个路由器转发时,数值减 1。当 HL 值减为 0 时,路由器丢弃该分组,向源节点发送 ICMPv6 “超时”报文。如果 HL 值大于 1,则将减 1 后的值写到 HL 字段中,进入下一个处理环节。

3）处理“载荷长度”字段

IPv6 的有效载荷长度等于“载荷长度”值，表示接收的有效载荷长度正确，那么路由器进入下一个处理环节；如果接收的有效载荷长度与“载荷长度”值不符，表示接收报文出错。路由器就丢弃这个报文，向源节点发送 ICMPv6“参数问题”报文。

4）处理“下一个报头”字段

如果 IPv6 的基本报头中“下一个报头”字段值为 0，表示该分组用扩展报头，并且是“逐跳选项”的扩展报头。按 IPv6 协议关于“逐跳选项”扩展报头的要求去处理。

5）进行路由选择

路由器使用目的 IPv6 地址和本地路由表中的内容进行比较，来确定转发接口和下一跳的 IPv6 地址，转发该报文。如果没有找到合适的路由，路由器就丢弃这个报文，向源节点发送 ICMPv6“目标不可到达”报文。

6）检查报文长度

确定下一跳端口之后，需要检查转发的报文长度是否大于端口链路的最大传输单元长度 MTU。如果大于端口链路的 MTU 值，路由器就丢弃该报文，向源节点发送 ICMPv6“报文过大”报文。如果小于端口链路的 MTU 值，路由器通过该端口转发报文。

需要注意的是，以上同样是给出了一个 IPv6 路由器在转发一个普通 IPv6 报文时简化了的处理过程。这里没有考虑以下几种情况。

（1）讨论中忽略了 IPv6 报文头中“通信类型”与“流标记”的处理细节。

（2）讨论中简化了 IPv6 报文头中“逐跳选项”等扩展报头的处理细节。

在研究了转发 IPv4 报文与 IPv6 报文的过程之后，我们会发现：转发一个 IPv6 数据包的过程比转发一个 IPv4 数据包要简单得多。因为转发一个 IPv6 报文时，报文头长度是固定的，不需要重新计算头校验和，也不需要执行拆分操作。因此，路由器对 IPv6 报文处理与转发的过程明显比 IPv4 简单，这样做的好处是可以减小 IPv6 路由器的转发延时，提高网络系统性能。

问题 5-12：如何认识 IPv4 与 IPv6 地址的区别？

IPv6 地址与 IPv4 地址的比较如表 5-3 所示。

表 5-3　IPv6 地址与 IPv4 地址的比较

比较的项目	IPv4	IPv6
长度	32 位	128 位
表示法	点分十进制	冒号分十六进制，带零压缩与双冒号简化表示
分类	分为 A、B、C、D、E 等 5 类地址	不按地址类型划分，而是按传输类型划分
网络地址表示	子网掩码或前缀长度	前缀长度
回送地址	127.0.0.0	::1
公网地址	单播地址	可汇聚全球单播地址
自动配置地址	169.254.0.0/16	链路本地地址 FE80::/64

续表

比较的项目	IPv4	IPv6
多播地址	224.0.0.0/4	FF00::/8
广播地址		未定义
未指定地址	0.0.0.0	::(0:0:0:0:0:0:0:0)
专用地址	10.0.0.0/8 172.16.0.0/12 192.168.0.0/16	FEC0::/48

问题5-13:术语辨析:子网掩码与前缀。

1. 子网掩码

理解术语“子网掩码”的含义需要注意以下几点。

(1) 子网掩码是由一台主机或路由器分配的地址位。通过子网掩码可以获得IP地址的网络地址与子网地址长度信息。子网掩码属于站点内部的局部问题。

(2) 子网掩码的长度与对应的IP地址长度相同,IPv4为32位,IPv6为128位。子网掩码中的1表示对应一个IP地址的网络/子网地址位,0表示对应的主机地址位。

(3) 子网掩码可以通过静态方式(典型的是路由器),也可以通过动态方式(如动态主机配置协议(DHCP))获取。

2. 前缀

为了帮助缓解IPv4地址的压力,分类寻址方案扩展了支持无类别域间路由CIDR的地址方案(见RFC4632)。这提供了一种方便的分配连续地址范围的方式。CIDR需要一个类似于子网掩码的掩码,有时也被称为CIDR掩码。CIDR掩码不再局限于一个站点,而对全球性路由系统都是可见的。因此,除了网络号之外,核心Internet路由器必须解释和处理的掩码称为网络前缀。它被用于IPv4和IPv6地址管理。如表5-4所示前缀的例子和它们相应的IPv4或IPv6地址范围。

表5-4 前缀的例子和它们相应的IPv4或IPv6地址范围

前缀	前缀(二进制)	地址范围
0.0.0.0/0	00000000 00000000 00000000 00000000	0.0.0.0~255.255.255.255
128.0.0.0/1	10000000 00000000 00000000 00000000	128.0.0.0~255.255.255.255
128.0.0.0/24	10000000 00000000 00000000 00000000	128.0.0.0~128.0.0.255
198.128.128.192/27	11000110 10000000 10000000 11000000	198.128.128.192~198.128.128.223
165.195.130.107/32	10100101 11000011 10000010 01101011	165.195.130.107
2001:db8::/32	0010000000000001 0000110110111000 0000000000000000 0000000000000000 0000000000000000 0000000000000000 0000000000000000 0000000000000000	2001:db8::~2001:db8:ffff:ffff

3. 前缀长度

子网掩码中用多个 1 和 0 表示的网络/子网地址位与主机地址位适用于计算机识别，而不适合于人的识别与记忆。例如，如果告诉我们 IP 地址为 128.32.1.16 的子网掩码 11111111 11111111 11111110 00000000，我们必须一个个数出 1 的个数，才能够知道对应于这个 IP 地址的网络/子网地址位长度为 23。

前缀长度适合于人的记忆。例如，将 128.32.1.16 的子网掩码 11111111 11111111 11111111 10000000 表示为 128.32.1.16/25，我们就很容易知道这个 IP 地址对应的网络/子网地址位长度为 25。前缀长度表示的是网络/子网地址位长度。

问题 5-14：如何理解地址聚合？

扩展支持无类别域间路由 CIDR 的地址方案有利于提高 IP 地址的利用率，但是带来了路由表项条目数量增大的问题。地址聚合或称为路由汇聚可以减少路由器之间路由选择信息的交换量，提高路由器工作效率。

一些技术可以显著地减少路由表条目数，同时保持在 Internet 中到达所有目的地址的最短路径。典型的方法是 20 世纪 70 年代末发表在分层路由 KK77 上 Kleinrock 和 Kamoun 的研究。他们发现，如果网络拓扑安排为一个树，并且用对这个网络拓扑“敏感”方式来分配地址，这样可使用一个很小的路由表保持到所有目的地的最短路径（如图 5-15 所示）。

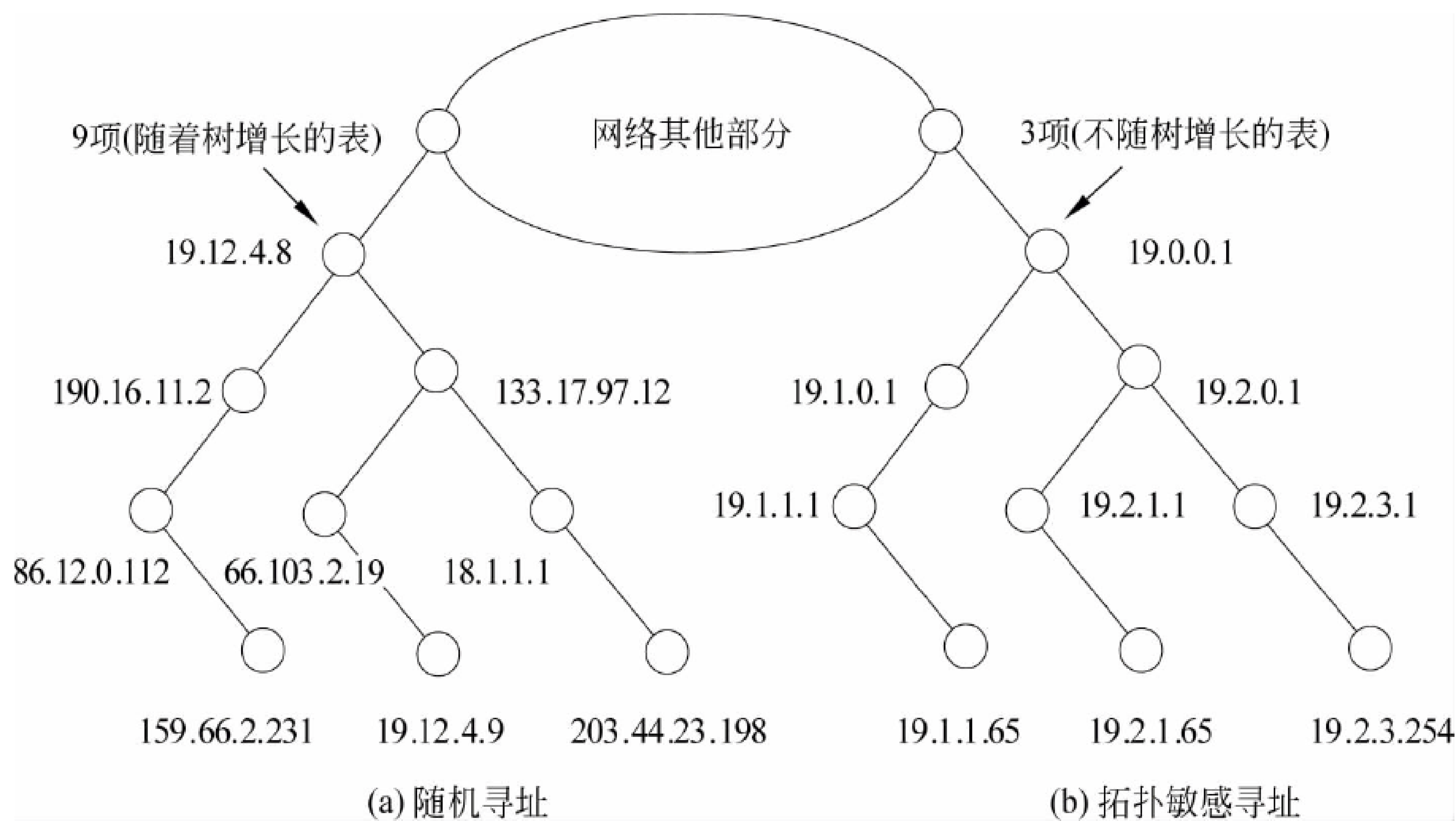


图 5-15　树状拓扑的网络结构示意图

在图论中一棵树的节点之间连接是没有循环的。对于一个网络中的路由器和链路构成的树，这就意味着树中任意两台路由器之间存在一个简单和不重复的路径。在树状拓扑的网络中，网络地址采用一个特殊方式分配，可以限制保存在路由器中路由条目的数量。

在图 5-15 中，圆表示路由器，线表示它们之间的通信链路。图的左侧（图 5-15(a)）和右侧（图 5-15(b)）显示的树状网络之间的区别是路由器的地址分配方式。

在图 5-15(a)中，节点地址基本上是随机的，路由器地址和它们之间的位置没有直接关系。在图 5-15(b)中，节点地址是根据路由器在树中位置关系分配的。从本例分析的每个

顶层路由器需要保存的路由条目的数量,可以看到一个很大的区别。

图5-15(a)树的根(顶级)路由器地址为19.12.4.8。为了知道每个可能的目的地址的下一跳节点地址,它需要一个在树中下层所有路由器的条目,如190.16.11.2、86.12.0.112等,包括自己共8条地址记录。图5-15(b)的根路由器地址为19.0.0.1,它的路由表中只要保存三条地址记录,这是由于所有路由器以前缀19.1和19.2开始。路由器19.0.0.1的路由表中只需将以19.1开始的目的地址显示下一跳为19.1.0.1。这种地址分配是递归的。研究表明,在Internet环境中,分层路由思想可用于以一种特定方式减少核心路由器中的路由条目数,这个过程称为路由聚合。

图5-16给出了路由聚合的示例。图5-16中左侧是三个地址前缀的聚合,其中,190.154.27.0/26与190.154.27.64/26数值相邻,可以被聚合。但是,前缀190.154.27.192/26不能在第一步被聚合,这是由于它不是数值相邻。当增加一个新前缀190.154.27.128/26(图中标有下画线),前缀190.154.27.192/26和190.154.27.128/26可以被聚合,形成前缀190.154.27.128/25。这个聚合与190.154.27.0/25相邻,因此它们可进一步被聚合成190.154.27.0/24。当增加前缀190.154.26.0/24(图中标有下画线),两个前缀可以进一步聚合为190.154.26.0/23。这样,原有的三个前缀和两个增加的前缀可聚合成一个前缀。

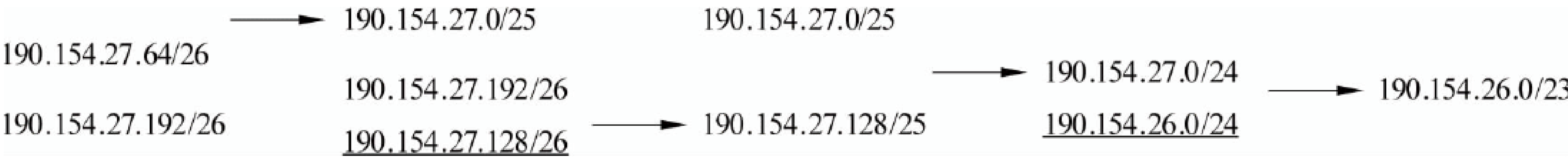


图5-16 路由聚合示例

显然,路由聚合可通过将相邻的多个IP前缀合并成一个短前缀,以覆盖更多地址空间。

问题5-15:IPv4与IPv6都定义了哪些特殊用途的地址?

IPv4和IPv6地址空间中都包括几个地址范围,不能用于单播地址分配,它们有特殊用途。RFC5735在2010年1月定义的特殊用途IPv4地址如表5-5所示。

表5-5 IPv4特殊用途地址

前　　缀	特殊用途	参考文献
0.0.0.0/8	本地网络中的主机,仅作为源IP地址使用	RFC1122
10.0.0.0/8	专用网络地址,不会出现在公共Internet上	RFC1918
127.0.0.0/8	Internet主机回送地址,通常只用127.0.0.1	RFC1122
169.254.0.0/16	“链路本地”地址,只用于一条链路,通常自动分配	RFC3927
172.16.0.0/12	专用网络地址,不会出现在公共Internet上	RFC1918
192.0.0.0/24	IETF协议分配(IANA保留)	RFC5736
192.0.2.0/24	批准用于文档中的TEST-NET-1地址,不会出现在公共Internet	RFC5737
192.88.99.0/24	用于6to4任播地址	RFC3068
192.168.0.0/16	专用网络地址,不会出现在公共Internet上	RFC1918
198.18.0.0/15	用于基准和性能测试	RFC2544

续表

前　缀	特 殊 用 途	参考文献
198.51.100.0/24	TEST-NET-2 地址，批准用于文档中	RFC5737
203.0.113.0/24	TEST-NET-3 地址，批准用于文档中	RFC5737
224.0.0.0/4	IPv4 组播地址(以前的 D 类)，仅作为目的 IP 地址使用	RFC5771
240.0.0.0/4	保留空间(以前的 E 类)，除了 255.255.255.255	RFC1112
255.255.255.255/32	本地网络(受限的)广播地址	RFC0919 RFC0922

RFC5156 在 2008 年 4 月定义的 IPv6 特定用途地址如表 5-6 所示。

表 5-6　IPv6 特殊用途地址

前　缀	特 殊 用 途	参考文献
::/0	默认路由条目，不用于寻址	RFC5156
::/128	未指定地址，可以作为源 IP 地址使用	RFC4291
::1/128	IPv6 主机回送地址，不用于发送出本地主机的分组中	RFC4291
::ffff:0:0/96	IPv4 映射地址，这种地址不出现在分组头部，只用于内部主机	RFC4291
::{ipv4-address}/96	IPv4 兼容地址，已过时，未使用	RFC4291
2001::/32	Teredo 地址	RFC4380
2001:10::/28	覆盖可路由的加密散列标识符，不出现在公共的 Internet 上	RFC4843
2001:db8::/32	用于文档和实例的地址范围，不出现在公共的 Internet 上	RFC3849
2002::/16	6to4 中继的 6to4 地址	RFC3056
3ffe::/16	用于 6Bone 实验，已过时，未使用	RFC3701
5f00::/16	用于 6Bone 实验，已过时，未使用	RFC3701
fc00::/7	唯一的本地单播地址，不用于全球性的 Internet 上	RFC4193
fe80::/10	链路本地单播地址	RFC4291
ff00::/8	IPv6 组播地址，仅作为目的 IP 地址使用	RFC4291

IPv4 和 IPv6 没有划定为特殊、组播或保留的地址范围都可供单播使用。一些单播地址空间(IPv4 的前缀 10/8、172.16/12 和 192.168/16 和 IPv6 的前缀 fc00::/7)被保留用于构建内部网络。来自这些范围的地址可用于一个站点或组织内部的主机和路由器之间通信，但不出现在全球 Internet 上。因此，这些地址有时也被称为不可路由的地址。

问题 5-16：如何理解 IP 地址中的子网广播地址与本地广播地址？

理解这个问题，需要注意以下几点。

1. 子网广播地址

每个 IPv4 子网中都保留着一个特殊地址作为子网广播地址。子网广播地址的主机字段中所有位为 1。如图 5-17 所示，IP 地址 128.32.3.12/23，那么只要根据子网掩码计算出主机号，将主机号各位取全 1，就可以得出子网广播地址 128.32.3.255。

需要注意的是：使用这种地址作为目的地的 IP 分组，也被称为定向广播。在理论上，这种广播可作为一个单个的 IP 分组通过 Internet 路由直至到达目的地子网，并以广播方式发送给子网中所有主机。但是，出于安全的考虑，目前 Internet 禁止使用定向广播。RFC919

IP地址	10000000 00100000 0000011 00001100	128.32.3.12
子网掩码	1111111111111111 11111110 00000000	255.255.254.0（/23）
主机地址	10000000 00100000 00000010 00000000	128.32.2.0
子网广播地址	10000000 00100000 00000011 11111111	128.32.3.255

图 5-17　子网广播地址的计算方法

IPv4 的各类广播，RFC1812 建议支持由路由器转发定向广播，它不仅可用，而且默认启用。但是 RFC2644 使这个策略发生逆转，默认路由器必须禁止转发定向广播分组。

2. 本地网络广播

除了子网广播地址，特殊用途的地址 255.255.255.255 被保留为本地网络广播，也称为有限广播，它根本不会被路由器转发。链路层的广播机制用于支持本地网络广播。广播地址通常与某些协议一起使用，如 UDP/IP 或 ICMP，因为这些协议不涉及 TCP/IP 那样的双方对话。IPv6 没有任何广播地址；广播地址可能被用于 IPv4 中，而 IPv6 仅使用组播地址。

问题 5-17：如何理解回送地址 127.0.0.0 的作用？

在正常的情况下，当一个 TCP/IP 应用发送一个数据字段时，它需要写上目的 IP 地址与源 IP 地址，并打成 IP 分组后，再提交数据链路层（如 Ethernet 协议）封装成帧，然后发送到网络中，由路由器逐跳转发直到目的主机。但是，IPv4 设计者为了测试实现 IP 协议的软件工作是否正常，设计了一个特殊的回送地址，范围是 127.0.0.1～127.255.255.255。如果要测试本地网络层执行 IP 协议的软件工作是否正常，可以由应用层执行一个"Ping 127.0.0.1"的命令，应用层的一个客户进程发出一个地址为 127.0.0.1 的 IP 分组，那么网络层软件将这个分组接收后转交给另一个客户进程，其过程如图 5-18 所示。因此，设计特殊的回送地址 127.0.0.1～127.255.255.255 的目的是用来测试主机高层与网络层软件的工作状态，确定本地进程之间的通信状态。

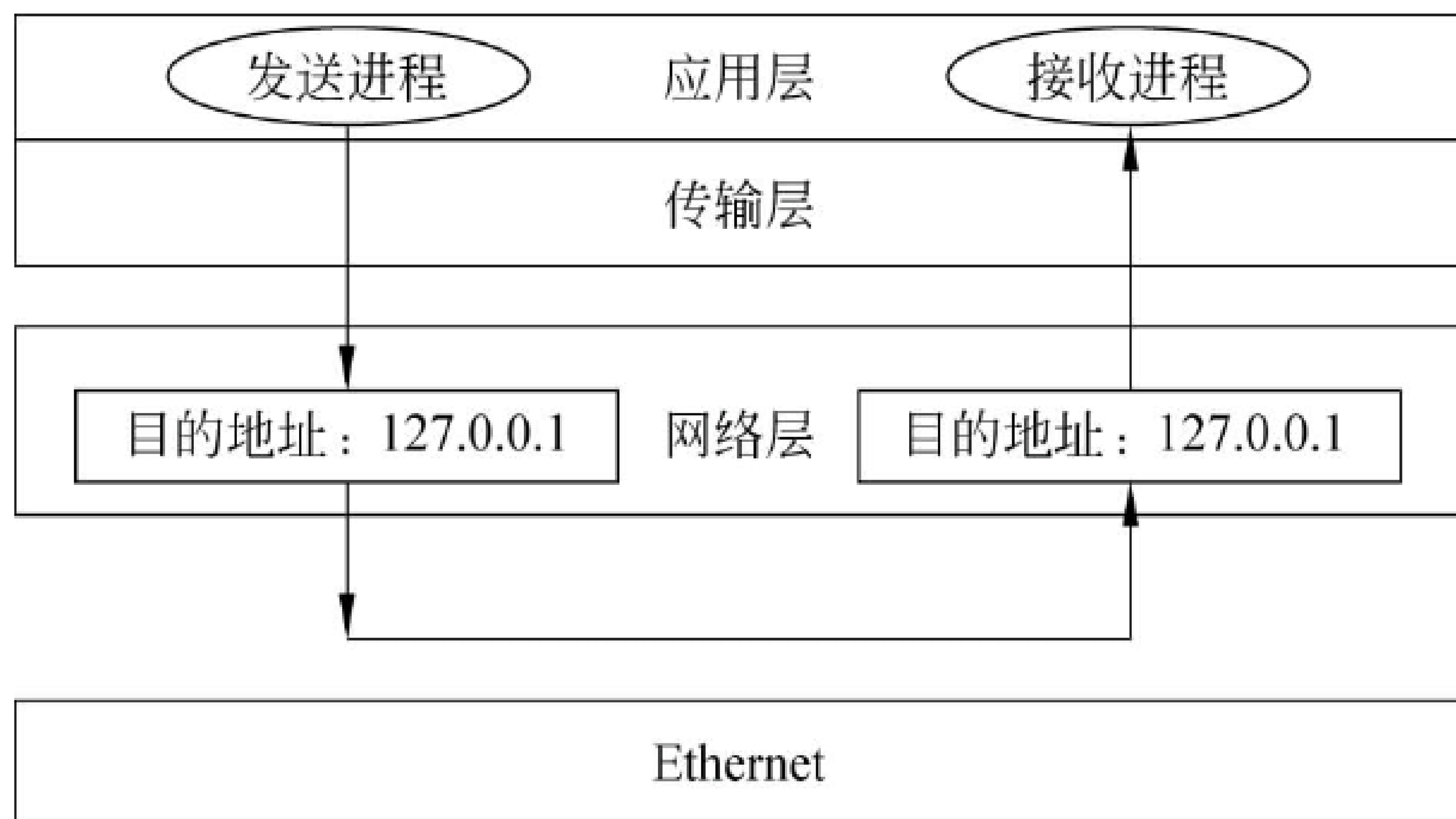


图 5-18　回送地址作用示意图

UNIX 系统（包括 Linux）实现回送检测软件的名称为 localhost。这样对于性能测试是有用的，例如，测量执行高层协议软件所需要用的时间。在 Linux 中，回送接口被称为 lo。

```
Linux%ifconfig lo
lo Link encap:Local Loopback
```

```
inet addr:127.0.0.1 Mask:255.0.0.0
inet6 addr: ::1/128 Scope:Host
UP LOOPBACK RUNNING MTU:16436 Metric:1
RX packets:458511 errors:0 dropped:0 overruns:0 frame:0
TX packets:458511 errors:0 dropped:0 overruns:0 carrier:0
collisions:0 txqueuelen:0
RX bytes:266049199(253.7 MiB)
TX bytes:266049199(253.7 MiB)
```

这里,我们看到本地回送接口的 IPv4 地址为 127.0.0.1,子网掩码为 255.0.0.0(对应于分级寻址中的 A 类网络号 127)。IPv6 地址::1 有一个 128 位的前缀,它表示只有一个地址。这个接口有一个 16KB 的 MTU(可配置为更大规模,最高可达 2GB)。从主机两个月前初始化开始,巨大流量(接近五十万个分组)已无差错地通过该接口。我们不希望在本地回送设备上看到错误,它实际上没有在任何网络上发送分组。

在 Windows 操作系统中,默认情况下没有安装 Microsoft 回送适配器,尽管它仍支持 IP 回送。当一个物理网络接口不可用时,可用于测试各种网络配置状况。对于 Windows Vista 或 Windows 7,在命令提示符下运行程序 hdwwiz,并手动添加 Microsoft 回送适配器。当它被执行时,ipconfig 命令显示如下(这个例子来自 Windows Vista)。

```
C:\> ipconfig /all
…
Ethernet adapter Local Area Connection 2:
    Connection-specific DNS Suffix:
    Description: Microsoft Loopback Adapter
    Physical Address: 02-00-4C-4F-4F-50
    DHCP Enabled: Yes
    Autoconfiguration Enabled: Yes
    Link-local IPv6 Address: fe80::9c0d:77a:52b8:39f0%18(Preferred)
    Autoconfiguration IPv4 Address: 169.254.57.240(Preferred)
    Subnet Mask: 255.255.0.0
    Default Gateway:
    DHCPv6 IAID: 302121036
    DNS Servers: fec0:0:0:ffff::1%1
    fec0:0:0:ffff::2%1
    fec0:0:0:ffff::3%1
NetBIOS over Tcpip: Enabled
```

这里,我们可看到该接口已被创建,已分配 IPv4 和 IPv6 地址,并显示为一系列虚拟以太网设备。现在,这台计算机具有以下回送地址。

```
C:\> ping 127.1.2.3
Pinging 127.1.2.3 with 32 bytes of data:
Reply from 127.1.2.3: bytes=32 time<1ms TTL=128
Reply from 127.1.2.3: bytes=32 time<1ms TTL=128
Reply from 127.1.2.3: bytes=32 time<1ms TTL=128
Reply from 127.1.2.3: bytes=32 time<1ms TTL=128
```

```
Ping statistics for 127.1.2.3:
Packets: Sent=4, Received=4, Lost=0(0%loss),
Approximate round trip times in milli-seconds:
Minimum=0ms, Maximum=0ms, Average=0ms
C:\>ping ::1
Pinging ::1 from ::1 with 32 bytes of data:
Reply from ::1: time<1ms
Reply from ::1: time<1ms
Reply from ::1: time<1ms
Reply from ::1: time<1ms
Ping statistics for ::1:
Packets: Sent=4, Received=4, Lost=0(0%loss),
Approximate round trip times in milli-seconds:
Minimum=0ms, Maximum=0ms, Average=0ms
C:\>ping 169.254.57.240
Pinging 169.254.57.240127.1.2.3 with 32 bytes of data:
Reply from 169.254.57.240: bytes=32 time<1ms TTL=128
Reply from 169.254.57.240: bytes=32 time<1ms TTL=128
Reply from 169.254.57.240: bytes=32 time<1ms TTL=128
Reply from 169.254.57.240: bytes=32 time<1ms TTL=128
Ping statistics for 169.254.57.240:
Packets: Sent=4, Received=4, Lost=0(0%loss),
Approximate round trip times in milli-seconds:
Minimum=0ms, Maximum=0ms, Average=0ms
```

问题 5-18：全 1 的 IP 地址与 host-ID 全 1 的 IP 地址区别是什么？

在 IPv4 协议的地址体系中，广播地址有两种：受限广播地址与直接广播地址。

(1) 全 1 的 IP 地址(255.255.255.255)为受限广播地址。当一个 IP 分组的目的 IP 地址为 255.255.255.255 时，路由器将不向外部互联的网络转发该分组，将其广播功能限制在本网络中。

(2) 在 A 类、B 类或 C 类 IP 地址中，host-ID 为全 1 的 IP 地址是直接广播地址。例如，一个主机连接在 191.11.0.0/16 的网络中，它发送一个目的地址为 191.11.255.255 的分组，显然目的地址的 host-ID 为全 1，这是一个直接广播地址。路由器将不向外转发该分组，该分组只能在网络号为 191.11.0.0 的网络中，以广播方式发送给所有主机。

受限广播地址、直接广播地址、"这个网络上的特定主机"地址、回送地址，以及保留地址都不会出现在 Internet 上。

问题 5-19：子网划分时 subnet-ID 能够取全 1 吗？

回答这个问题需要注意以下几点。

(1) 在 1985 年 8 月公布的 RFC950 文档规定中，子网号不能是全 0 或全 1。做出这种规定的主要理由是如果将全 0 或全 1 的子网号分配给子网后，在向一个网络的所有子网进行多播时可能会引起混乱。

(2) 但是在 1995 年 12 月公布的 RFC1878 中，对 RFC950 提出了不同的看法，允许将全

0或全1的子网号分配给子网。因为无类别域间路由CIDR中是用前缀去表示地址块的范围,子网号为全0和全1的地址可以分配给子网,但是主机号全0和全1的广播地址不能分配给主机。实际地址规划中常常会涉及这个问题。

(3) 需要注意的是:RFC1878并不是Internet标准,而是属于“提供信息的”文档。因此,在使用时,需要注意路由器制造商是不是支持RFC1878。

问题5-20:如何理解网络专用地址的作用?

如果一个组织需要组建一个专用网络,不准备连接到Internet,但是网络需要运行TCP/IP,那么有三种方法可供选择。

(1) 可以和连接到Internet的用户一样,向Internet管理部门申请一个唯一的IP地址,但不与Internet实现物理的连接。这种方法的好处是:一旦该组织决定要连接到Internet,所有主机和路由器可以非常容易地接入到Internet。但是如果这个组织并没有打算连接到Internet,那么这个IP地址资源就有可能被浪费掉了。

(2) 可以正常申请和使用A类、B类或C类IP地址,但是不需要到Internet管理部门去注册。因为网络并没有连接到Internet上,它是孤立的,因此也就可以不考虑IP地址的唯一性问题。这种方法的好处是处理起来简单,但是缺点比较明显。用户可能错误地认为他所使用的就是正常的Internet的唯一的IP地址。在某一天要接入Internet时会造成混乱。

(3) 为了克服以上方法的缺点,Internet管理机构在分配IP地址时,就预留了专用网络使用的地址。预留地址如表5-7所示。

表5-7　专用网络使用的地址

类	网　络　号	总数
A	10.0.0.0～10.255.255.255	1
B	172.16.0.0～172.31.255.255	16
C	192.68.0.0～192.68.255.255	256

由于表5-7所示的地址是Internet管理机构预留给专用网络使用的,因此任何组织使用都不需要向Internet管理机构申请,所有网络管理人员都应该知道这些地址是为专用网络内部使用的。这类地址在专用网络内部是唯一的,但是在Internet中并不是唯一的。目前,很多电子政务专网、企业专网都是采用如10.0.0.0/8中的部分地址块。

问题5-21:如何认识IPv6地址的特点?

可以从以下几个方面来认识IPv6地址的特点。

1. IPv6地址表示方法

1) 基本表示方法

2006年,RFC4291“IPv6 Addressing Achitecture”对IPv6地址空间结构与地址基本表示方法进行了定义,定义了冒号十六进制表示法。

(1) IPv6的128位地址按每16位划分为一个位段。

(2) 每个位段被转换为一个4位的十六进制数。

(3) 位段之间用冒号隔开。

可以按以下步骤形成一个冒号十六进制IPv6地址。

第一步:用二进制格式表示一个IPv6地址。

0010000000000001000000000000000000000000000000000000000000000000

0000001010101010000000000000111111111110000010001001110001011010

第二步:将这个128位的地址按每16位划分为8个位段。

0010000000000001 0000000000000000 0000000000000000 0000000000000000

0000001010101010 0000000000001111 1111111000001000 1001110001011010

第三步:将每个位段转换成十六进制数,并用冒号隔开,结果应该是

2001:0000:0000:0000:02AA:000F:FE08:9C5A

那么得到的一个冒号十六进制IPv6地址与最初给出的一个用128位二进制数表示的IPv6地址是等效的。由于十六进制和二进制之间的进制转换,比十进制和二进制之间的进制转换更容易,因此IPv6的地址表示法采用十六进制数。每一位十六进制数对应4位二进制数。128位的IPv6的地址实在太长,人们很难记忆。在IPv6网络中,节点的IPv6地址都是自动配置的。

2)零压缩法

一个IPv6地址中可能会出现多个二进制数0,可以规定一种方法,通过压缩某个位段中的前导0,可以进一步简化IPv6地址的表示。例如,“00D3”可以简写为“D3”;“02AA”可以简写为“2AA”。但是“FE08”就不可以简写为“FE8”。如果出现“000A”可以简写为“A”。同时需要注意的是,每个位段至少应该有一个数字,“0000”可以简写为“0”。

前面给出了一个IPv6地址的例子:

2001:0000:0000:0000:02AA:000F:FE08:9C5A

根据前导零压缩法,上面的地址可以进一步简化表示为

2001:0:0:0:2AA:F:FE08:9C5A

有些类型的IPv6地址中包含一长串0。为了进一步简化IP地址表达,在一个以冒号十六进制表示法表示的IPv6地址中,如果几个连续位段的值都为0,那么这些0就可以简写为::,称为双冒号表示法。

那么,前面的结果又可以简写为:2001::2AA:F:FE08:9C5A。

同样根据零压缩法,链路本地地址FE80:0:0:0:0:FE:FE9A:4CA2可以简写为FE80::FE:FE9A:4CA2。组播地址FF02:0:0:0:0:0:0:2可以简写为FF02::2。

需要注意的问题有以下几点。

(1)在使用零压缩法时,不能把一个位段内部的有效0也压缩掉。例如,不能将FF02:30:0:0:0:0:0:5简写为FF2:3::5,而应该简写为FF02:30::5。

(2)::双冒号在一个地址中只能出现一次。例如地址0:0:0:2AA:12:0:0:0,一种简化的表示法是::2AA:12:0:0:0,另一种表示法是0:0:0:2AA:12::,不能把它表示为::2AA:12::。

(3)要确定::之间代表了被压缩的多少位0,可以数一下地址中还有多少个位段,然后用8减去这个数,再将结果乘以16。例如,在地址FF02:3::5中有三个位段(FF02、3和5),可以根据公式计算:(8−3)×16=80,则::表示有80位的二进制数字0被压缩。

3）IPv6 前缀

理解 IPv6 前缀(Format Prefix，FP)的概念，需要注意以下几个问题。

(1) IPv6 不支持子网掩码，只支持前缀长度表示法。

(2) IPv6 地址类似于 IPv4 的 CIDR，IPv6 前缀可以表示为"地址/前缀长度"。

如果一个节点的 IPv6 地址为 2001:FA2:0:FE08::9C5A；地址前缀长度为 48。那么：

(1) 节点的子网号为 2001:FA2::/48。

(2) 同时表示节点地址与前缀时写为 2001:FA2:0:FE08::9C5A/48。

(3) 前缀 48 位表示地址的前 48 位为网络地址，之后的 80 位可以分配给网络中的主机，可以分配给主机的地址数量共有 2^{80} 个。

2. IPv6 地址分类与特点

1）IPv6 地址的分类

根据 2006 年 RFC4291 对 IPv6 地址的分类，IPv6 地址分为：单播地址、组播地址、任播地址三种基本类型(其结构如图 5-19 所示)。

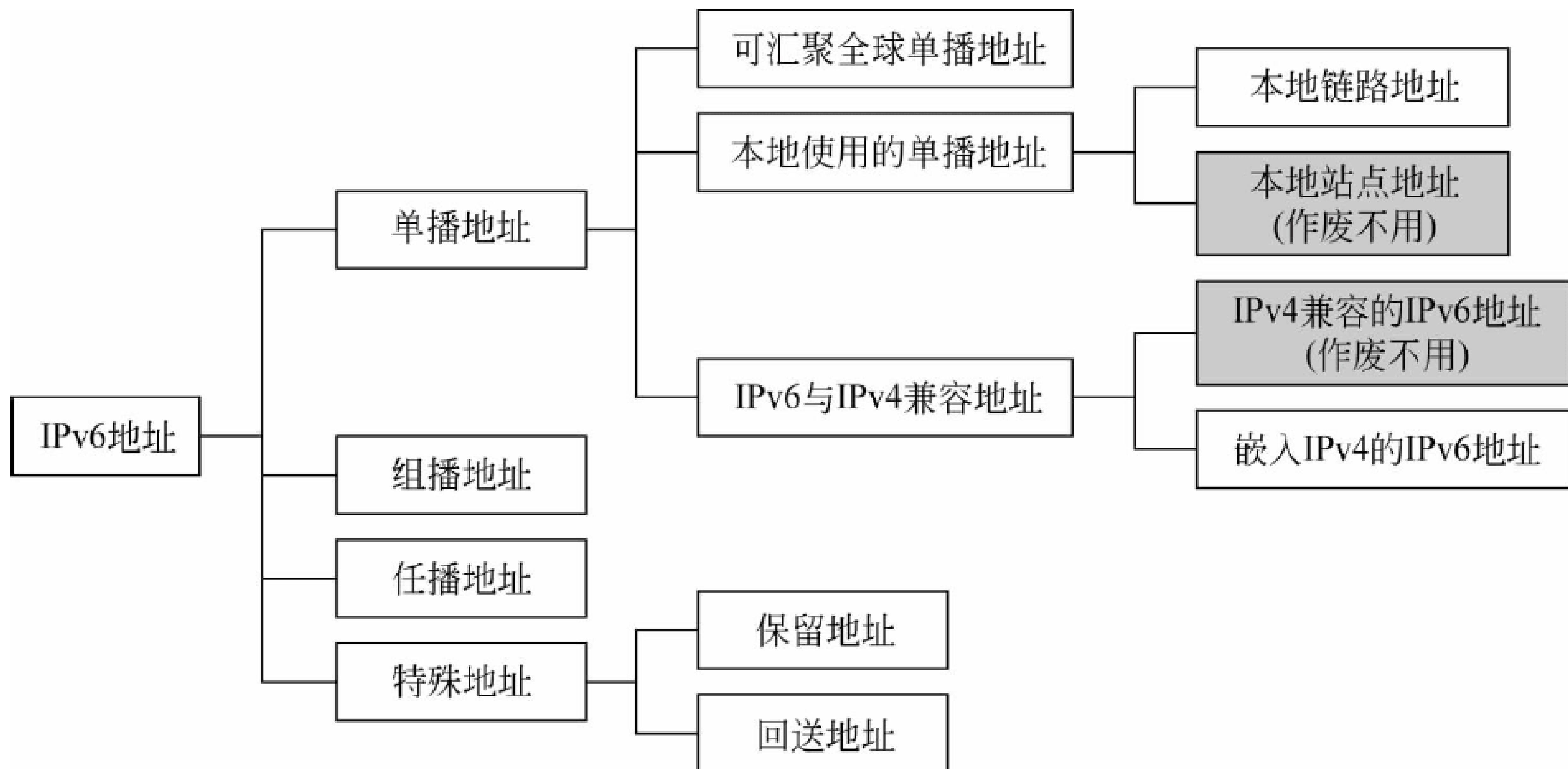


图 5-19　IPv6 地址类型

需要注意的是：2006 年发布的 RFC4291 废止了 2003 年发布的 RFC3513，RFC3513 废止了 1998 年发布的 RFC2373。由于 IPv6 地址一直处于改进的过程中，因此在发现问题之后，一定会有新的 RFC 修改发布，来修订地址分配方案。例如，本地站点地址(Site-Local Address)类似于 IPv4 中的专用地址，本地站点地址出现在公网上，有可能带来一定的安全问题，因此被废除。目前，单播地址中已经不包括本地站点地址。因此，研究 IPv6 地址需要密切注意新的 RFC 文档发布的信息，并且关于 IPv6 地址问题的 RFC 文档有很多，更新很快。

2）IPv6 地址分类特点

RFC4291 对 IPv6 地址分类做出了以下描述。

(1) IPv6 地址分为：单播地址、组播地址、任播地址三种基本的类型。

(2) 单播地址用来标识路由器、主机的某一个接口。发往单播地址的报文，被传送给该地址标识的网络接口。

(3) 组播地址用来标识一组属于不同节点的网络接口。发送到多播地址的报文，被交

付给由该地址标识的所有接口。

(4) 任播地址用来标识一组属于不同路由器的接口,发送到任播地址的报文,被交付给由该地址标识的一组接口中距离“最近”的一个。

(5) IPv6 不使用广播地址,广播地址的功能由组播地址代替。

IPv6 地址表示举例:

(1) 单播地址 2001:DB8:0:0:0:08AB:200C:0002 可以写为 2001:DB8::8AB:200C:2。

(2) 多播地址 FF01:0:0:0:0:0:0:101 可以写为 FF01::101。

(3) 回送地址 0:0:0:0:0:0:0:1 可以写为::1。

(4) 未指定地址 0:0:0:0:0:0:0:0 可以写为::。

理解 RFC4291 对 IPv6 地址分类特点的描述,需要注意以下几个问题。

(1) IPv6 地址中未指定的保留地址为 0:0:0:0:0:0:0:0(或::),它与 IPv4 中地址 0.0.0.0 相同。这个地址不分配给任何节点的接口。只有在一种情况:当初始化主机发送 IPv6 报文时,主机不知道自己的地址,它就在自己的源地址中写入保留地址 0:0:0:0:0:0:0:0。

(2) IPv6 的回送地址 0:0:0:0:0:0:0:1 与 IPv4 的回送地址 127.0.0.0 意义、使用方法相同。回送地址不能出现在任何网络中,主机和路由器不能为该地址广播任何寻址信息。

“Ping”应用程序可以发送一个将回送地址作为目的地址的报文,以测试 IP 软件能否接收或发送一个报文。一个客户进程可以用回送地址发送一个报文给本机的另一个进程,用来测试本地进程之间的通信状况。

问题 5-22:如何认识 IPv6 单播地址的特点?

单播 IPv6 地址可以分为:

(1) 可汇聚全球单播地址;

(2) 链路本地地址;

(3) 嵌有 IPv4 的 IPv6 地址。

可汇聚全球单播地址可以用于全球网络中寻址,而链路本地地址的寻址范围是受限制的。早期规定的站点本地地址已经被废止。

1. 可汇聚全球单播地址

IPv4 地址结构的一个主要缺点是按照地址类型去划分,IPv4 地址结构与节点的地理位置无关。由于 Internet 并不主张有国家边界,IPv4 地址是直接分配给最终用户的,即使后期能力使 IPv4 地址由国家或地区来分配,这种状态也没有很大的变化。理论上来说,一个法国的站点既可以连接法国或欧洲的 ISP,也可以与美国的 ISP 连接。这个站点既可能有一部分法国 ISP 分配的地址,也可能有一部分美国 ISP 分配的地址。随着电信市场自由化的发展,这种情况变得越来越复杂。而 IPv6 地址改变了地址分配策略,它将基于最终用户的分配思路改变为基于 ISP 的分配思路。

IPv6 地址设计的一个重要特点是:可以有效地支持多级寻址和路由。可汇聚全球单播地址是 IPv6 的公网地址,类似于 IPv4 的单播地址。具有可汇聚全球单播地址的报文,可以在全球范围内 IPv6 网络中有效地路由和转发。

根据 RFC1881 文档的规定,IPv6 地址由 IANA 与地区性组织分配与管理。目前,参与 IPv6 地址分配与管理的 5 个地区性组织是:IP 研究欧洲网络合作中心(RIPE-NCC)、北美

Internet 网络信息中心(INTERNIC)、亚太网络信息中心(APNIC)、拉丁美洲网络信息中心(LACNIC)与非洲网络信息中心(AfriNIC)。

可汇聚全球单播地址的分配过程如下。

(1) 全球 IPv6 地址分配组织 ICANN 首先通过设置顶级汇聚标识符(Top-Level Aggregation Identifier,TLA ID),将地址分块;不同的 TLAID 地址块,分配给不同的地区性 IP 地址管理组织。

(2) 获得 TLA ID 地址块的地区性组织继续将地址块分成多个下一级汇聚标识符(Next-Level Aggregation Identifier,NLA ID)地址块,不同的 NLA ID 地址块,分配给申请的国家 IP 地址管理组织,或连接在主干网上的大型 ISP。

(3) 国家 IP 地址管理组织或大型 ISP 再将 NLA ID 地址块细分成站点级汇聚标识符(Site-Level Aggregation Identifier,SLA ID)地址块,分配给不同的站点。

(4) 站点继续将 SLA ID 地址块分块,分配给它的子网。子网获得带有 TLA ID-NLA ID-SLA ID 结构的地址后,与子网的主机、路由器的接口标识符(Interface ID),组成网络节点的 IPv6 可汇聚全球单播地址。

可汇聚全球单播地址层次结构如图 5-20 所示。

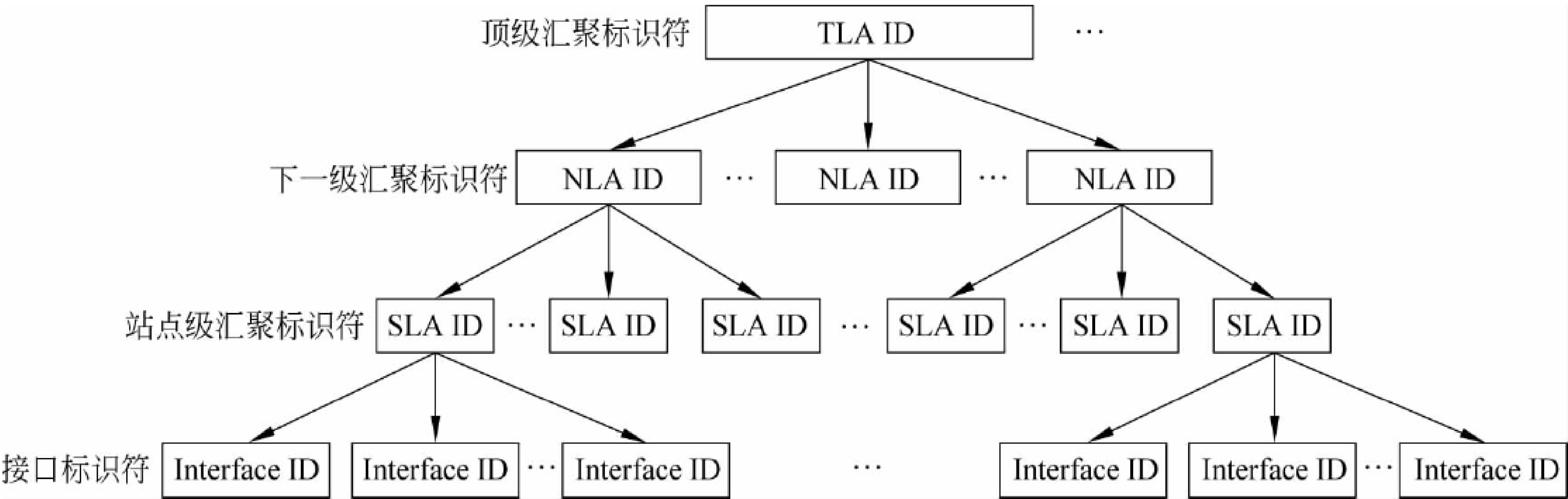


图 5-20　可汇聚全球单播地址层次结构

从 IPv6 可汇聚全球单播地址分配过程可以看出:任何一个网络节点的 IPv6 地址都可以聚合成具有带有 TLA ID-NLA ID-SLA ID 结构特征的路由。可汇聚全球单播地址结构可以支持三层的网络拓扑结构。

(1) 第一层:由顶级汇聚标识符 TLA ID 与下一级汇聚标识符 NLA ID 描述的是公网拓扑,它反映出多级的 ISP 的网络结构。

(2) 第二层:站点拓扑反映出站点内部的网络拓扑结构。

(3) 第三层:接口 ID 标识唯一地标识出节点的网络接口。

这种按 ISP 层次结构、有序分配的单播地址,可以清晰地反映互联网络的层次结构。这样的地址结构既有利于路由器快速地选择路由,又能够有效地控制骨干网的路由规模。RFC2450 推荐的可汇聚全球单播地址具体的分配策略如图 5-21 所示。

被推荐的可汇聚全球单播地址结构包括以下几个字段。

(1) 格式前缀(FP)。

可汇聚全球单播地址的格式前缀 FP 长度为 3b,数值为 001。



图 5-21 可汇聚全球单播地址分配策略

(2) 顶级汇聚标识符(TLA ID)。

顶级汇聚标识符 TLA ID 由 IANA 分配给指定的注册机构。顶级汇聚标识符在路由层次结构中是最高级。由于 TLA ID 的长度是 13b,因此 IPv6 网络中最多有 2^{13}=8192 个不同的顶级 TLAID。

(3) 保留(Reserved for Future Use,RES)。

保留字段长度为 8b,留作将来扩展 TLA 和 NLA 字段时用。目前保留位必须置 0。

(4) 下一级汇聚标识符(NLA ID)。

下一级汇聚标识符 NLA ID 长度为 24b,由申请 TLAID 地址块的机构分配。NLA ID 能够按照它们的分级寻址结构。TLA ID、NLA ID 的结构反映了公共网络的网络拓扑与路由。

(5) 站点级汇聚标识符(SLA ID)。

站点级汇聚标识符 SLA ID 由机构分配给它下属的子网。SLA ID 长度为 16 位,可以支持最多支持 65 535 个子网。

(6) 接口标识符(Interface Identifier,INTERFACE ID)。

接口 ID 标识特定一个子网节点的网络接口。设置单播地址接口 ID 长度为固定的 64b,其目的是为了便于将常用的 Ethernet 的 48 位 MAC 地址映射为 EUI-64 地址。

IANA 负责 IPv6 地址空间的分配,并且委派 5 个地区组织执行地址分配与管理的任务。目前,IANA 已分配的可汇聚全球单播地址空间如表 5-8 所示。

表 5-8 已分配的可汇聚全球单播地址空间

前缀(十六进制)	前缀(二进制)	分配说明
2001::/16	0010 0000 0000 0001	北美、欧洲、亚太、拉美、非洲等地区
2002::/16	0010 0000 0000 0010	IPv6 到 IPv4 转换机制
2003::/16	0010 0000 0000 0011	欧洲地区
2400:0000::/19 2400:2000::/19 2400:4000::/21	0010 0100 0000 0000 000 0010 0100 0000 0000 001 0010 0100 0000 0000 0100 0	亚太地区
2600:0000::/22 2604:0000::/22 2608:0000::/22 260C:0000::/22	0010 0110 0000 0000 0000 00 0010 0110 0000 0010 0000 00 0010 0110 0000 1000 0000 00 0010 0110 0000 1100 0000 00	亚太地区

续表

前缀(十六进制)	前缀(二进制)	分配说明
2A00:0000::/21 2A01:0000::/23	0010 1010 0000 0000 0000 0 0010 1010 0000 0001 0000 000	欧洲地区
3FFE::/16	0010 1111 1111 1110	6Bone

2. 链路本地地址

理解链路本地地址需要注意以下几个问题。

(1) 链路本地地址用格式前缀“1111111010”标识。

(2) 链路本地地址用于同一链路上的相邻节点之间的通信。链路本地地址的作用范围是本地链路,即以路由器为界的单个链路范围内。

(3) 路由器不转发带有链路本地地址的报文。

(4) 链路本地地址是自动配置的。

链路本地地址结构如图 5-22 所示。链路本地地址是由特定的一个前缀与接口 ID 组成的。

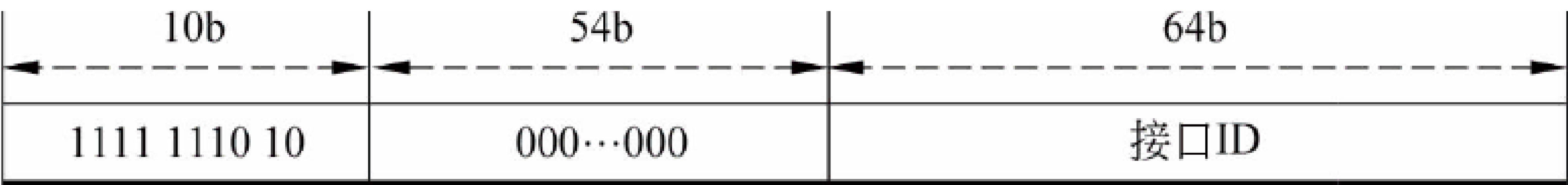


图 5-22　链路本地地址结构

前缀又是由 10 位的格式前缀“1111 1110 10”与 54 位的 0 组成。如果考虑一台 Ethernet 网卡的 MAC 地址为 08-00-02-12-2A-1B 的 PC,那么这台 PC 的链路本地地址的前缀与子网部分的 64 位二进制比特序列可以写为

1111 1110 10000000 00000000 00000000 00000000 00000000 00000000 00000000

用十六进制表示为

FE80:0000:0000:0000

将 MAC 地址映射成 EUI-64 之后,形成的链路本地地址为

FE80:0000:0000:0000:0000:0811:0212:2A1B

那么,这台 PC 的链路本地地址可以写为“FE80::800:212:2A1B”。

设计链路本地地址的好处表现在:如果一个小型的 IPv6 网络,将几台 PC 互联起来,在没有路由器的情况下,PC 可以快速、自动地生成本地链路 IPv6 地址进行通信。链路本地地址还用于邻居发现协议、自动地址配置等单一链路寻址的应用中。

3. 嵌有 IPv4 的 IPv6 地址

在 IPv6 与 IPv4 共存的情况下,还应该有一种能够嵌有 IPv4 的 IPv6 地址。RFC4213 对这种处于 IPv4 向 IPv6 过渡阶段的特殊的地址形式做了规定。它可以在主机和路由器在 IPv4 路由基础设施的结构中,动态地通过隧道方式传输 IPv6 报文。为了达到这个目的,研究人员曾经定义了两种形式的 IPv6 单播地址结构:IPv4 兼容的 IPv6 地址、嵌有 IPv4 的 IPv6 地址。

IPv4 映射地址又称为嵌有 IPv4 的 IPv6 地址,其结构如图 5-23 所示。

RFC4213 文档已经明确废止了 IPv4 兼容的 IPv6 地址与自动隧道机制。

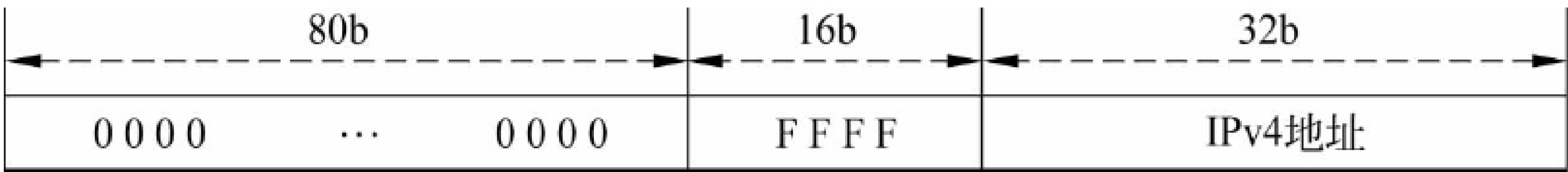


图 5-23　嵌有 IPv4 的 IPv6 地址结构

问题 5-23：如何认识 IPv6 组播地址的特点？

理解 IPv6 组播地址需要注意以下几个问题。

1. IPv6 组播地址的结构

组播也称作多播，它是指一个源节点发送一个报文能够被一个组中多个目的节点所接收。IPv6 组播数据流的运行方式与 IPv4 基本上是相同的。任意位置上的 IPv6 节点，都可以侦听到带有 IPv6 组播地址的组播报文。IPv6 节点可以随时加入或离开一个组播组。

2. 组播地址中字段的意义

IPv6 组播地址的结构如图 5-24 所示。

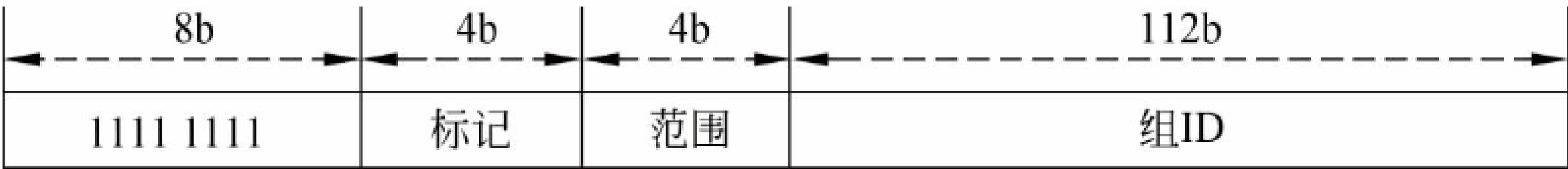


图 5-24　IPv6 组播地址的结构

IPv6 组播地址的格式前缀是“1111 1111”，因此一个 IPv6 组播地址总是以“FF”开始。除了格式前缀，组播地址中的字段有标记、范围与组 ID。

1）标记

该字段的长度是 4b，是组播地址标记。在 RFC4291 中定义为 | 0 | R | P | T |，其中最高位必须为 0；R 位表示是否为内嵌汇聚点地址的多播地址；P 位表示是否为基于单播网络前缀的多播地址；T 为暂态位标记，当 T=0 时，表示当前的组播地址是由 IANA 所分配的一个永久分配的、众所周知的组播地址。T=1 则表示当前的组播地址是一个临时组播地址。

2）范围

该字段的长度是 4b，表示组播报文在 IPv6 网中发送的范围。RFC4291 定义的范围字段值与它表示的范围如表 5-9 所示。

表 5-9　范围字段值意义

范围字段值	表示的范围	范围字段值	表示的范围
0	预留	8	组织本地范围
1	节点本地范围	E	全球范围
2	链路本地范围	F	预留
5	站点本地范围		

例如，对于永久分配的组播地址标记 T=0，如果是链路本地范围则范围字段值的最低位值为 2，那么组播地址前缀的前 16 位应该是“1111 1111 0000 0010”，那么它可以记为“FF02”。

3）组 ID

该字段用来标识组播组，它的长度为 112b，可以生成 2^{112} 个组 ID。目前，RFC2373 并没有将所有的 112b 都用于定义组 ID，建议使用 112 位中的低 32b 定义组 ID，而其余 80b 置 0。目前使用的组地址结构如图 5-25 所示。

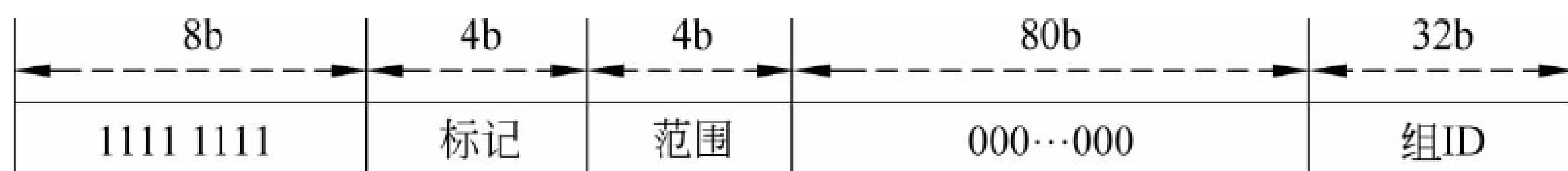


图 5-25　目前使用的组播地址结构

从 FF01∷到 FF0F∷的组播地址是保留的专用地址，用于标识本地节点，以及本地链路范围内的所有节点。IPv6 也有一些特殊的组播地址。

长度为 112b 的组 ID 可以标识给定范围的多播组。这个多播组可以是永久的，也可以是临时的。例如，需要为网络时间协议（Network Time Protocol，NTP）服务器组永久地分配组 ID 为 43（十六进制）。那么，下面所有的地址都属于组 43，但是地址覆盖的范围不同。

（1）FF01∷43 意思是：与发送方同一个节点上的 NTP 服务器。

（2）FF02∷43 意思是：与发送方同一个链路上的 NTP 服务器。

（3）FF05∷43 意思是：与发送方同一个站点上的 NTP 服务器。

（4）FF0E∷43 意思是：与发送方同一个网络上的 NTP 服务器。

要得到永久分配的 IPv6 组播地址的最新列表，可以访问：

http://www.ian6.org/assignments/IPv6-multicast-addresses

问题 5-24：如何认识 IPv6 任播地址的特点？

理解 IPv6 任播地址需要注意以下几个问题。

1．任播地址的基本概念

组播地址用于一个节点对多个节点通信，而任播地址则用于一个节点对多个节点中的一个节点的通信。“任播地址”也称为“泛播地址”或“任意点传送地址”。带有任播地址的分组将被路由器转发给与其连接在同一个网络中，按路由协议计算“最近”距离的一个路由器接口。

2．任播地址的特点

RFC4291 定义了子网-路由器任播地址。它是由一个给定接口的子网前缀来创建的，地址中的其余位都设为 0。子网-路由器任播地址的结构如图 5-26 所示。

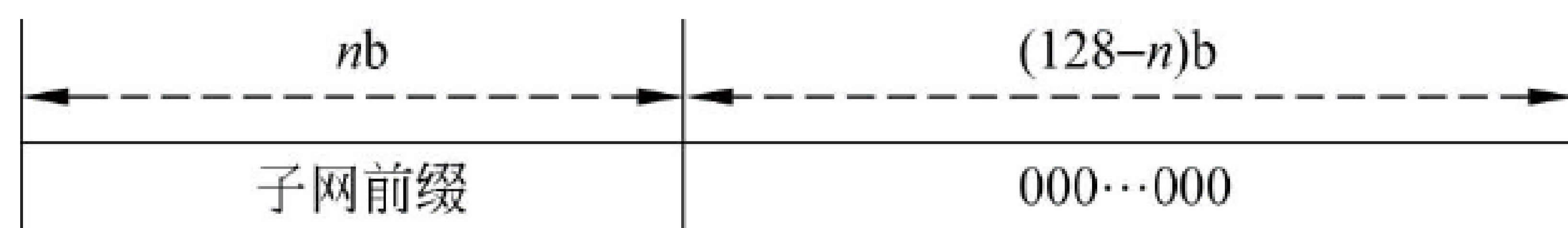


图 5-26　子网-路由器任播地址的结构

任播地址是一种新型的地址，具有广泛的应用前景。典型的应用是在移动 IPv6 的动态家乡代理地址发现机制中。但是从安全角度考虑，目前任播地址只能作为目的地址，分配给路由器，不配置给主机。所有连接到一个子网的路由器接口，都有这个子网的子网-路由器任播地址，用于同连接到该子网中最近的路由器之间的信息交互与通信。任播地址与单播地址结构相同，将一个单播地址分配给多个网络接口时，这个单播地址就成了任播地址。分

配任播地址时必须说明。

问题 5-25：如何认识 IPv6 主机地址与路由器地址的不同？

理解这个问题需要注意以下几个问题。

1. 主机 IPv6 地址

在 IPv4 网中，如果一台主机通过一块网卡接入网络，那么只需要给这台主机的该接口分配一个 IPv4 地址。但是，由于 IPv6 地址是按传输类型分类的，因此一台主机同样是通过一块网卡接入 IPv6 网络时需要给网卡接口分配多个类型的 IPv6 地址。

一台 IPv6 主机应该具有的 IPv6 单播地址是：

(1) 接口的链路本地地址；

(2) 接口的全球单播地址；

(3) 回送地址。

因此，从一个主机接口的地址情况来看，IPv6 主机至少拥有两个地址，一个是可以用于本地链路通信的链路本地地址，另一个是可以在全网路由的全球单播地址。

2. 主机需要随时侦听的地址类型

一台主机的网络接口需要随时侦听以下目的地址的 IPv6 报文。

(1) 本地节点范围内所有节点的组播地址(FF01::1)的报文。

(2) 本地链路范围内所有节点的组播地址(FF02::1)的报文。

(3) 以本节点的单播地址为目的地址的报文。

(4) 同组组播地址的报文。

3. 路由器 IPv6 地址

一台路由器的每个网络接口应该具有的 IPv6 地址是：

(1) 链路本地地址；

(2) 全球单播地址；

(3) 回送地址；

(4) 子网-路由器任播地址。

4. IPv6 路由器的接口随时侦听的地址类型

路由器的每个网络接口需要随时侦听以下目的地址为组播地址的 IPv6 报文是：

(1) 节点本地范围内所有节点的组播地址(FF01::1)的报文；

(2) 节点本地范围内所有路由器的组播地址(FF01::2)的报文；

(3) 本地链路范围内所有节点的组播地址(FF02::1)的报文；

(4) 本地链路范围内所有路由器的组播地址(FF02::2)的报文；

(5) 站点本地范围内所有路由器的组播地址(FF05::2)的报文；

(6) 全球单播地址的报文；

(7) 同组组播地址的报文。

问题 5-26：如何认识 ICMPv6 协议的特点？

IP 协议本身并没有为主机提供直接的方法来发现发往目的地址的 IP 数据包是否失败。另外，IP 也没有提供直接的方式来获取诊断信息。例如，沿途的路由器也不估计分组传输的时间。为了弥补这些不足而设计了 Internet 控制报文协议 ICMP(RFC0792 与

RFC4443)，与 IP 协议结合起来使用，提供与 IP 协议配置和 IP 数据分组处置相关的诊断和控制信息。ICMP 通常被认为是 IP 层本身的一部分，它需要依赖 IP 协议来传输。因此，ICMP 既不是一个网络层协议，也不是一个传输层协议，而是位于两者之间的协议。认识 ICMPv6 协议的特点，需要将它与 ICMPv4 协议做一个比较。

1. ICMPv4 协议的特点

1）ICMPv4 协议研究的背景

IP 协议提供的是尽力而为的服务。IP 协议的优点是简洁，缺点是缺少差错控制和查询机制。IP 报文一旦发送出去，是否到达目的主机，以及在传输过程中出现哪些错误，源主机的 IP 协议是不知道的。在这种情况下，如果出现一些问题，例如路由器找不到可以到的目的网络，报文生存时间超过而必须被丢弃，以及目的主机在规定的时间内不能接收属于同一个报文的所有分片该怎么办。因此，必须通过一种差错报告与查询、控制机制来了解信息，决定如何处理。ICMPv4 协议就是为解决以上问题而设计的，它是配合 IPv4 协议使用的。ICMPv4 的差错与查询、控制功能对于保证 IPv4 协议的可靠运行是至关重要的。

2）ICMPv4 协议的特点

ICMPv4 协议的特点主要表现在以下几个方面。

（1）ICMPv4 本身是网络层的一个协议，但是它的报文不是直接传送给数据链路层，而是要封装成 IPv4 报文，然后再传送给数据链路层。

（2）从协议体系上看，ICMPv4 只是要解决 IPv4 协议可能出现的不可靠问题，不能独立于 IPv4 协议而单独存在，它是 IPv4 协议的一个组成部分。

（3）ICMPv4 设计的初衷是用于 IPv4 协议在执行过程中的出错报告，实际上是由路由器向源主机报告传输出错的类型，差错处理与控制需要由高层协议完成。

2. ICMPv6 协议

1）ICMPv6 协议特点

IPv6 结构体系中同样设计了 ICMPv6（RFC4443），它的主要功能也是进行错误报告和网络诊断等。和 IPv4 一样，ICMPv6 是 IPv6 协议的一个组成部分。ICMPv6 具备了 ICMPv4 的所有基本功能，不同之处主要有两点，一是它删除了一些不再使用的过时报文类型，定义了其他一些新的功能与报文；二是 ICMPv6 合并了 ICMP、IGMP 与 ARP、RARP 等多个协议的功能（如图 5-27 所示）。

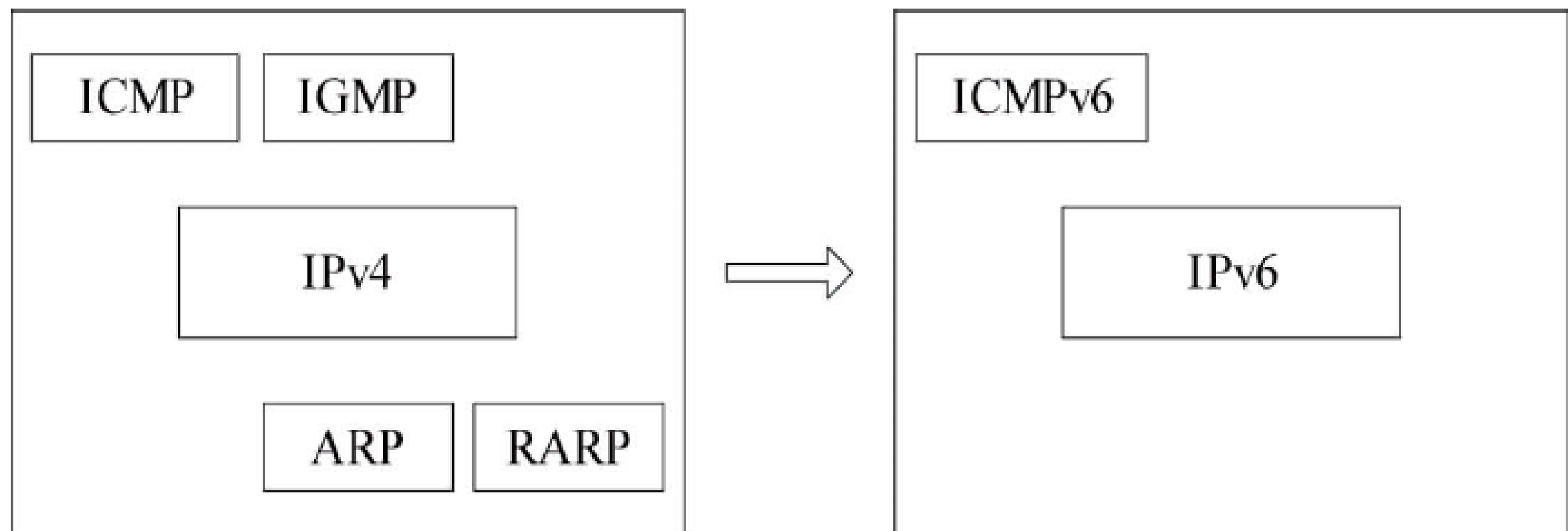


图 5-27　ICMPv4 与 ICMPv6 的区别

ICMPv6 的控制信息类型可主要划分为两种：差错报文与信息报文。

差错报文主要用于报告 IPv6 报文在传输过程中出现的错误。常用的 ICMPv6 差错报文类型主要有：

(1) 目的不可达;

(2) 报文过大;

(3) 超时与参数问题。

信息报文主要用于提供网络诊断功能与附加的主机功能。常用的ICMPv6信息报文类型主要有:组播侦听发现与邻节点发现。

2) ICMPv6报文与IPv6报文的关系

ICMPv6所有报文都是封装在IPv6报文中来传送的,ICMPv6报文与IPv6报文的关系如图5-28所示。每一个ICMPv6报文在传送时都必须附加上一个IPv6基本报文,如果有扩展报头,还需要加上一个或多个扩展报头。在离它最近的扩展报头中的“下一个报头”值应该为58。因此,对于IPv6报文来说,它是将ICMPv6报文作为一般的报文来处理的,只是从扩展报头中的“下一个报头”值是不是为58,来判断它发送的是不是ICMPv6报文。

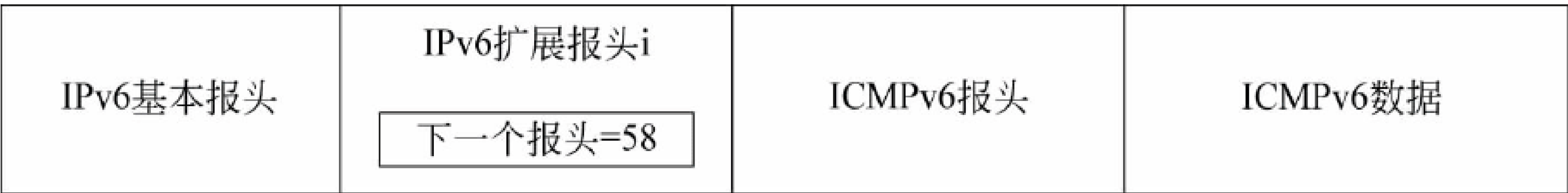


图5-28 ICMPv6报文与IPv6报文的关系

3) ICMPv6报文的结构

ICMPv6报文结构如图5-29所示。

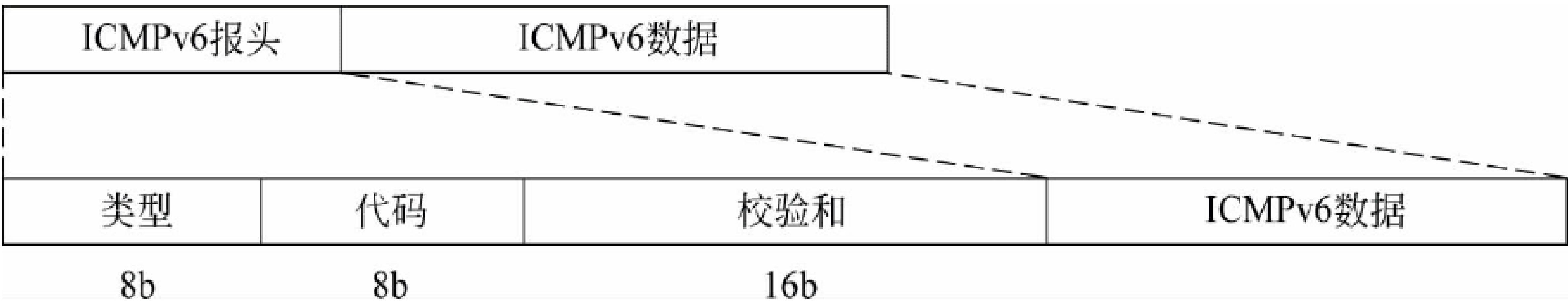


图5-29 ICMPv6报文结构

ICMPv6报头由类型、代码与校验和三个部分组成。

类型字段长度为8b,它用来表示ICMPv6的报文类型。代码字段长度也是8b,它是从属于类型字段,可以在基本的类型上再细分出新的类型。ICMPv6报文的校验和长度为16b,它计算的对象是ICMPv6类型字段加上IPv6的伪报头。

表5-10给出了常用的报文类型与对应的类型字段值。类型字段值在0~127为差错报文,值在128~225为信息报文。

表5-10 类型字段值与对应的报文类型

ICMPv6报文类	类型字段值	所对应的报文类型
差错报文	1	目的不可到达(destination unreachable)
	2	报文过大(packed too big)
	3	超时(time exceeded)
	4	参数问题(paramenter problem)

续表

ICMPv6 报文类	类型字段值	所对应的报文类型
信息报文	128	回声请求(echo request)
	129	回声应答(echo reply)
	130	组成员查询(group membership query)
	131	组成员报告(group membership report)
	132	组成员减少(group membership reduction)
	133	路由器请求(router solicitation)
	134	路由器公告(router advertisement)
	135	邻节点请求(neighbor solicitation)
	136	邻节点公告(neighbor advertisement)
	137	重定向(redirect)

ICMPv6 报文分为差错报文和信息报文等两类。ICMPv6 的报文类型,其结构如图 5-30 所示。

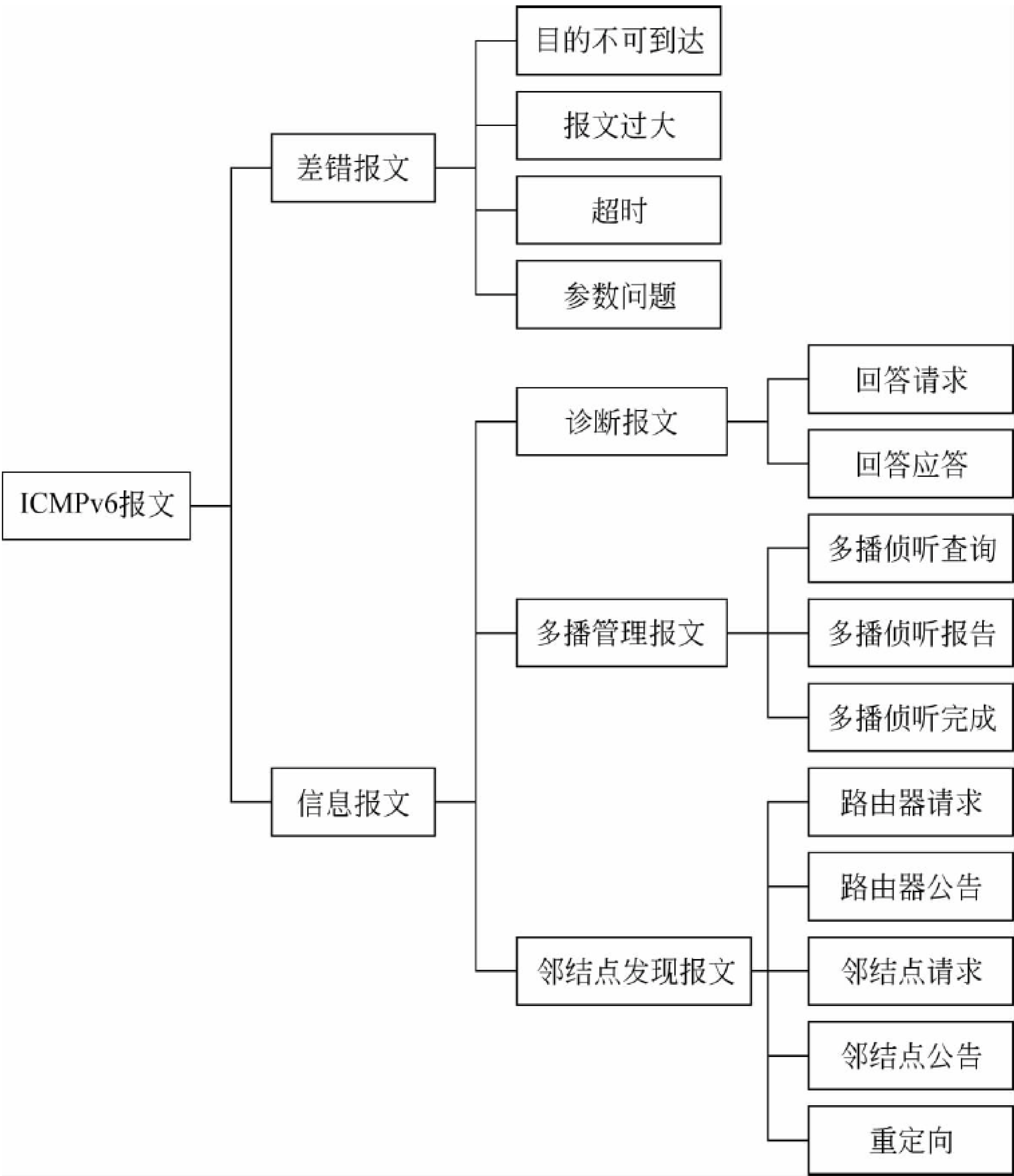


图 5-30 ICMPv6 的报文类型与结构

3. ICMPv6 差错报文

ICMPv6 差错报文主要包括4种基本类型:目的不可到达、报文过大、超时与参数问题。

1)目的不可到达

表5-11给出了4种情况导致该报文不能到达目的节点和目的端口,这时路由器或主机将向源节点发送“目的不可到达”报文,报告IPv6报文在传输过程中出现差错。

表5-11 代码字段值与目的不可到达原因

代码字段值	对应的意义
0	没有到达目的节点的路由,路由器无法转发
1	路由器或防火墙禁止与某个目的节点通信
2	未指定
3	因无法解析到目的节点链路层MAC地址,导致目的地址不可到达
4	IPv6报文已经传送到目的IP节点,但是不能递交给目的TCP或UDP端口

2)报文过大

当路由器发转发一个报文时,发现报文的长度大于准备转发该报文的下一跳出口链路MTU,那么路由器只能丢弃该报文,并且向发送该报文的源节点发送“报文过大”报文,报告IPv6报文在传输过程中出现差错。

3)超时

当路由器接收到一个报文时发现报文的跳数限制字段值为0或1,路由器将丢弃该报文,并且向发送该报文的源节点发送“超时”报文,报告IPv6报文在传输过程中出现差错。

4)参数问题

当路由器或主机接收到的一个报文的基本报头或扩展报头出现错误,而不能继续处理时,路由器将丢弃该报文,并且向发送该报文的源节点发送“参数问题报文”,报告IPv6报文在传输时出差错的类型。

4. ICMPv6 信息报文

ICMPv6 信息报文主要包括三种基本类型:诊断报文、多播组管理报文与邻节点发现报文。

1)ICMPv6 诊断报文

ICMPv6 诊断报文结构如图5-31所示。

8b	8b	16b
类型 (128/129)	代码	校验和
标识		序列号
数据		

图5-31 诊断报文结构

在IPv6网络中,任何一个节点在接收到回送请求报文后,一定要发送回送应答报文作为答复。为了便于实现诊断功能,一个节点一般都要为ICMPv6回送请求报文与回送应答

报文提供应用层接口。回送请求报文的类型字段值为128，回送应答报文的类型字段值为129。

诊断报文的回送请求报文与回送应答报文，它是用来实现ping与tracert等功能，以检查目的地址是否能够到达。可以通过ping与tracert功能的实现，加深对ICMPv6诊断报文工作原理的理解。诊断报文还可以用于“路径MTU发现”。

ping是测试目的主机是否能够到达的一种通用的方法。图5-32给出了一台主机ping另一台主机的结构示意图。

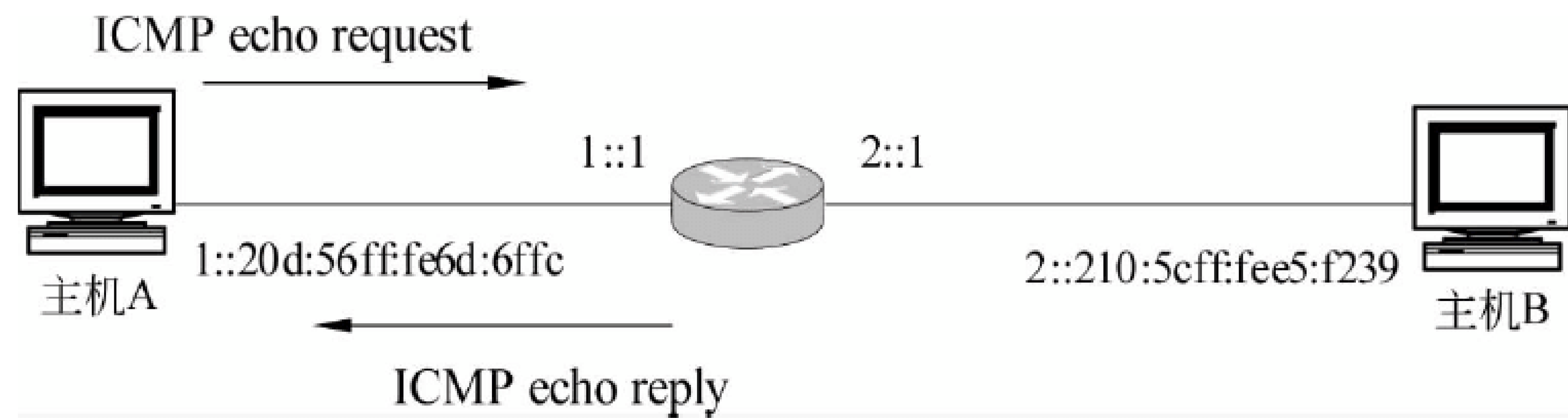


图5-32　一台主机ping另一台主机的结构示意图

主机A要ping主机B时，它可以使用以下结构的ICMP echo request报文。下面给出一个用Network Monitor捕获的简化的ICMPv6 echo request的报文。

```
Internet Protocol v6:
    Version= 6 (IPv6)
    traffic class= 0x0 (0)
    flow label= 0x00000 (0)
    playload length= 0x0028 (40)
    next header= 0x003A (58,ICMPv6)
    hop Limit= 128
    source address= 1::20d:56ff:fe6d:6ffc
    destination address= 2::210:5cff:fee5:f239
Internet Control Message Protocol v6:
    type= 128 (echo request)
    code= 0 (0x0)
    checksum= 0x76dc (correct)
    id= 8 (0x0008)
    data (32 bytes)
```

主机B在接收到echo request报文之后，发出下面的echo reply应答报文。

```
Internet Protocol 6:
    Version= 6 (IPv6)
    traffic class= 0x00 (0)
    flow label= 0x00000 (0)
    playload length= 0x0028 (40)
    next header= 0x003A (58,ICMPv6)
    hop Limit= 63
    source address= 2::210:5cff:fee5:f239
    destination address= 1::20d:56ff:fe6d:6ffc
```

```
Internet Control Message Protocol v6:
    type=129(echo reply)
    code=0(0x0)
    checksum=0x75dc(correct)
    id=0(0x0)
    sequence number=24(0x18)
    data(32 bytes)
```

显然,主机 A 在接收到 echo reply 报文之后,就可以做出主机 A 与主机 B 可以通信的结论。因此,ping 是本地主机检查它是否能够与另外一台主机通信的主要工具,也是实现域名解析的主要方法。很多操作系统都有 ping 工具。例如,在安装了 IPv6 的 Windows. NET Server 中,输入 ping 命令后,系统执行 ping 命令会显示以下结果。

```
F:\ ping 2::210:5cff:fee5:f239%1
ping 2::210:5cff:fee5:f239%1 from 1::20d:56ff:fe6d:6ffc
with 32 bytes of data
reply from 2::210:5cff:fee5:f239 : time<1ms
reply from 2::210:5cff:fee5:f239 : time<1ms
reply from 2::210:5cff:fee5:f239 : time<1ms
reply from 2::210:5cff:fee5:f239 : time<1ms
ping statistics from 1::20d:56ff:fe6d:6ffc
pacet: sent=4,received=4,lost=0(0%loss)
approximate round trip time in milli-seconds:
minimum=0 ms,maximum=0 ms,average=0 ms
```

从以上结果中可以看出,在 Windows. NET Server 中执行 ping 命令,除了能够检查主机之间的连通性之外,还可以知道报文传输的往返时间,同时它还可以进行 IPv6 的地址解析。

tracert 是 IP 网络中重要的诊断工具之一,它可以给出到达目的地址的路径。tracert 工作原理如图 5-33 所示。

为了获得从源主机 A 到目的主机 B 的路径,启动 tracert 程序,步骤如下。

第一,发送一个跳数限制值为 1 的 echo request ICMP 报文给目的节点。第一个接收到的路由器将跳数限制值 1 减 1 为 0 的报文丢弃,并向源节点发送一个超时 ICMP 报文。那么,源节点就得到了第一个路由器的地址。

第二,发送一个跳数限制值为 2 的 ICMP echo request 报文给目的节点。第二个接收到的路由器也会因为跳数限制值的原因,丢弃报文,并向源节点发送一个超时 ICMP 报文。那么,源节点就得到了第二个路由器的地址。

第三,继续执行以上的过程,直至 ICMP echo request 报文到达目的节点,目的节点发送回一个 ICMP echo reply 应答报文。这样,源节点就可以获得一个完整的从源节点到达目的节点的路径列表。

Windows. NET Server 操作系统支持 IPv6 的 tracert 功能。在 Windows. NET Server 上执行 tracert 命令之后,可以得到以下结果。

```
F:\>tracert fec0::f282:2b0:d0ff:fee9:4143%1
```

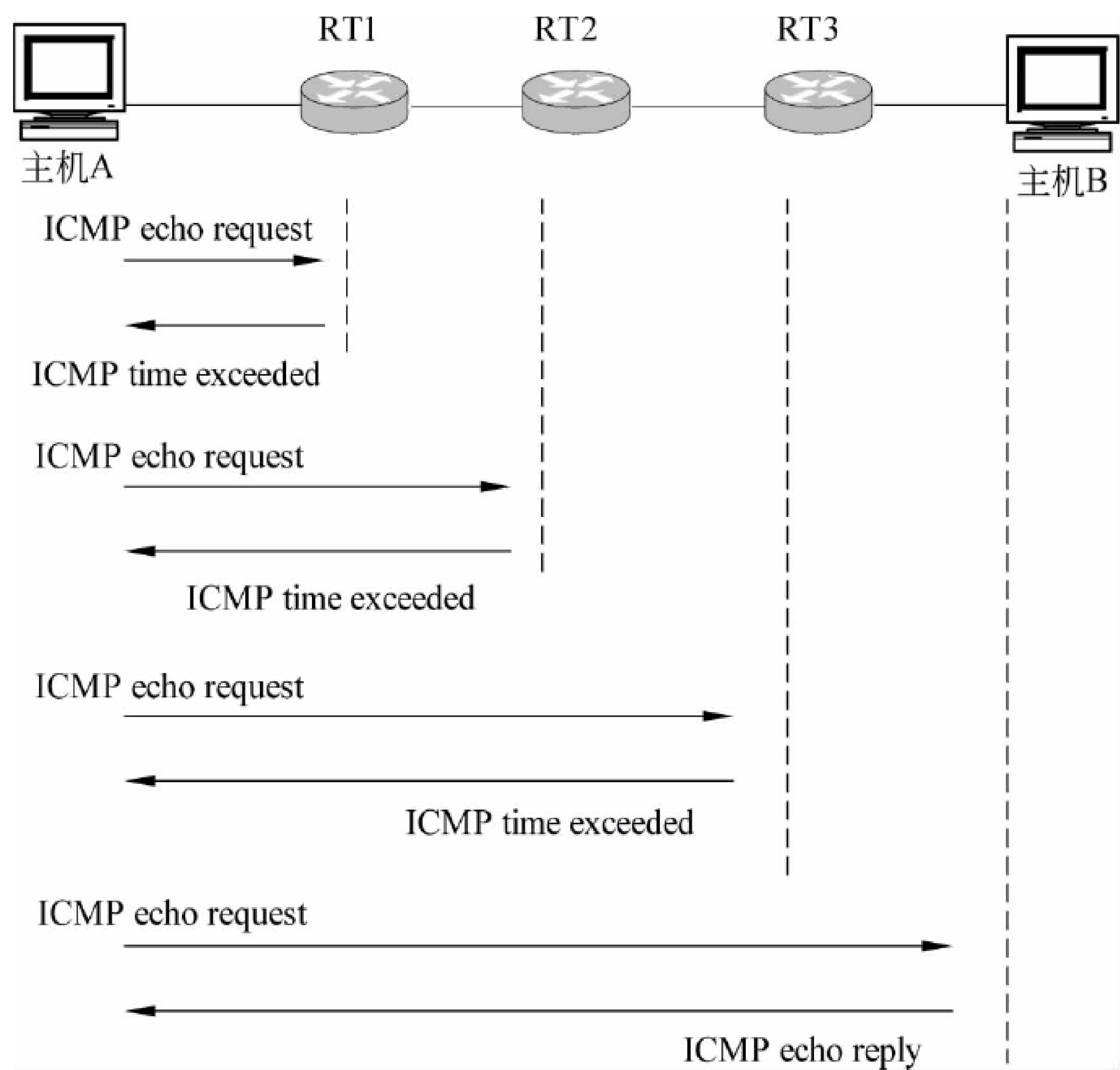


图 5-33　tracert 工作原理示意图

```
tracert route to fec0::f282:2b0:d0ff:fee9:4143%1 over a maximumof 30 hops
1  < 1ms  < 1ms  < 1ms   fec0::f241:2b0:d0ff:fead:243d
2  < 1ms  < 1ms  < 1ms   fec0::f2ac:2b0:d0ff:fea5:d347
3  < 1ms  < 1ms  < 1ms   fec0::f282:2b0:d0ff:fee9:4143
tracert complete.
```

2）多播组管理报文与多播侦听发现协议 MLD

IPv4 的多播组管理是通过 IGMP 实现的。IPv6 采用多播侦听发现(Multicast Listener Discovery,MLD)协议与 ICMPv6 多播组管理报文实现了对多播组的管理。RFC2710 对多播侦听发现 MLD 协议进行了描述和定义。

IPv6 将具有特定多播地址的多台主机的集合称为多播组。多播组成员的身份是动态的,一个多播组中的成员数是没有限制的,主机可以在任何时候加入或离开一个多播组。不是多播组成员的主机,也可以向一个多播地址发送多播通信流。多播组可以跨越多个 IPv6 路由器,即跨越多个子网。这种配置就要求 IPv6 路由器支持 IPv6 多播,同时也要求主机具有通过 MLD 协议来进行加入或撤除多播组的能力。

多播侦听发现 MLD 协议是 IPv6 路由器使用的一种协议,用以发现在路由器直接连接的网络中,是否有希望接收多播报文的节点(这些节点称为“多播侦听者”),以及多播侦听者对哪些组播地址感兴趣;根据 MLD 发现的信息,路由器使用多播路由协议,将多播报文转发到本地链路中存在的多播侦听者。

MLD 协议是使用 ICMPv6 多播组管理报文来实现的。MLD 报文的一般格式如图 5-34 所示。MLD 报文的 IPv6 报头中第一个“下一个报头”值为 0,表示紧接着 IPv6 基本报头之后的是“逐跳选项报头”中“路由器告警选项”;“逐跳选项报头”中“下一个报头”值为 58,表示紧接着的是 ICMPv6 多播组管理报文中的“MLD 报文”。

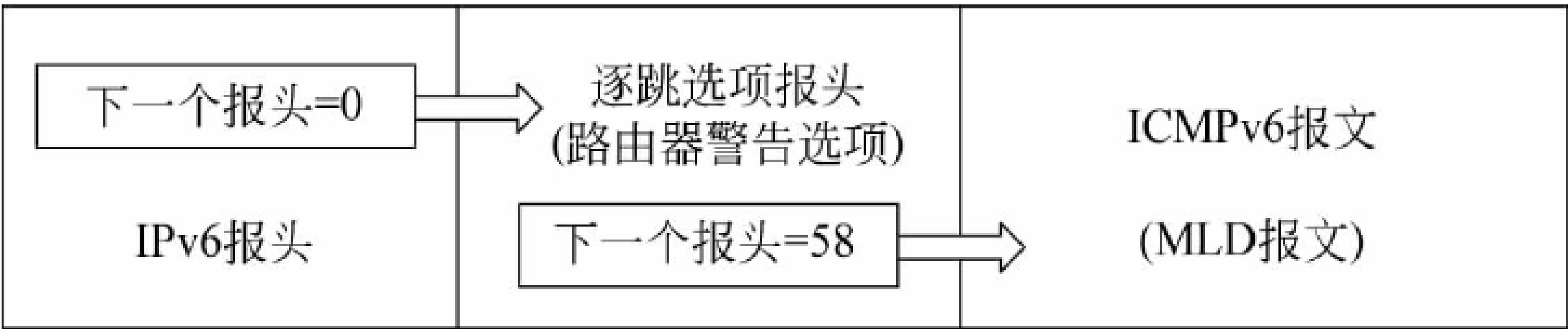


图 5-34　MLD 报文的一般格式

RFC2710 文档中定义了以下三种类型的报文。

(1) 查询报文(ICMPv6 报头类型编码为 130);

(2) 报告报文(ICMPv6 报头类型编码为 131);

(3) 已完成报文(ICMPv6 报头类型编码为 132)。

其中,查询报文又分为一般查询与特定多播地址查询。一般查询只是要知道哪个多播地址有侦听者,而特定多播地址查询是要知道该多播地址有多少个侦听者。MLD 报文跳数限制值为 1,发送时使用的目的地址分为不同的情况(如表 5-12 所示)。

表 5-12　MLD 报文使用的目的地址

MLD 报文类型	IPv6 目的地址
一般查询	本地链路范围内的所有节点(FF02::1)
特定多播地址查询	查询的多播地址
报告	报告的多播地址
已完成	本地链路范围内的所有节点(FF02::2)

RFC3810 定义了 MLD 协议第 2 版(MLDv2)。MLDv2 与第 1 版 MLDv1 兼容,增加了对特定源组播和非特定源组播地址过滤的功能。多播路由器通过多播路由协议,收集多播侦听者的信息,并将侦听状态通告给与它相邻的多播路由器。

3) 邻节点发现报文

邻节点发现(Neighbor Discovery,ND)是指:用一组 ICMPv6 信息报文来确定邻节点之间关系的过程。IPv6 的 ND 协议取代了 IPv4 的地址解析协议 ARP、ICMPv4 的路由器发现协议与 ICMPv4 重定向协议。

IPv6 的邻节点发现协议包括以下基本功能。

(1) 路由器发现;

(2) 前缀发现;

(3) 参数发现;

(4) 地址自动配置;

(5) 地址解析;

(6) 下一跳选择;

(7) 邻节点不可到达检测;

(8) 重复地址检测;

(9) 重定向。

邻节点发现报文主要有 5 种:路由器请求报文、路由器公告报文、邻节点请求报文、邻节点公告报文与重定向报文。

2007年9月公布的RFC4861取代了最初对邻节点发现(ND)协议进行了定义的RFC 2461。

问题5-27:如何理解IPv6地址自动配置功能?

理解IPv6地址的自动配置的概念,需要注意以下几个问题。

1. 地址配置的基本概念

在了解IPv6地址基本结构的基础上,需要讨论如何在路由器与主机上配置IPv6地址。主机接入IPv4网络时,需要采用手工的方式为它配置一个32位IPv4地址、子网掩码、默认网关地址,以及DNS地址。IPv6地址的自动配置功能可以实现即插即用的入网方式,减轻了网络管理员的很多工作负荷。

理解IPv6地址自动配置时,需要注意以下几个问题。

(1) IPv6协议定义了两种方法:无状态地址自动配置与有状态地址自动配置。

(2) 在不特别关注主机使用的确切IP地址,只要求该地址在全网是唯一的,并且能通过适当的方式进行路由选择时,就可以使用无状态地址自动配置。当主机对IP地址分配要求严格时,就应该使用有状态地址自动配置。在有状态地址自动配置中,DHCP服务器维护一个数据库,记录被分配的地址,以及主机的配置信息,主机可以从DHCP服务器中获得地址及其他配置信息。

(3) 在路由器公告中,无状态与有状态地址自动配置的处理是相互独立的,主机可以同时使用无状态与有状态地址自动配置。主机可以使用无状态地址自动配置来配置自己的地址,但是需要使用有状态地址自动配置来获取其他的参数与信息。

2. IPv6无状态地址自动配置

无状态是相对于有状态而言的。RFC1971、RFC2462文档定义了IPv6无状态地址自动配置(Stateless Address Autoconfiguration)的概念与配置过程。无状态地址自动配置的过程如下。

(1) 用主机网卡MAC地址生成的EUI-64接口标识符与链路地址前缀1111 111010(FF80::/64),自动生成一个本地链路地址。

(2) 主机发送邻居请求报文。进行地址重复检测,确定临时本地链路地址的唯一性。如果收到邻居通告报文,则说明已有节点在使用该本地链路地址,应停止使用;如果没有收到邻居通告报文,则说明该地址是唯一的,可以使用。

(3) 主机发送路由器请求报文,请求本链路上的路由器响应路由器通告报文。路由器通告报文包含各种路由信息与主机配置所需要的信息,如链路前缀、链路MTU、跳步限制、特定路由,以及是否使用地址自动配置、地址优先级、地址的有效期等。默认情况下,主机最多可以发送三个路由器请求报文,同时路由器也在周期性地发送路由器通告报文。

(4) 主机接收到路由器通告报文,主机将根据报文内容来设置跳步限制、链路MTU、重发定时器等参数。

(5) 如果存在前缀信息选项,要进行相应的标志位处理。

① 若链路标志为1,将报文中的前缀添加到前缀列表中。

② 若自治标志为1,用前缀和EIU-64接口标识符生成一个临时地址;通过重复地址检测来确定地址的唯一性。

(6) 如果“路由器公告”报文的管理地址配置标志位置1,则用有状态地址自动配置协议

获取其他的地址。

(7)如果"路由器公告"报文的有状态配置标志位置1,则用有状态地址自动配置协议获取其他的配置参数。

IPv6无状态地址自动配置的流程如图5-35所示。

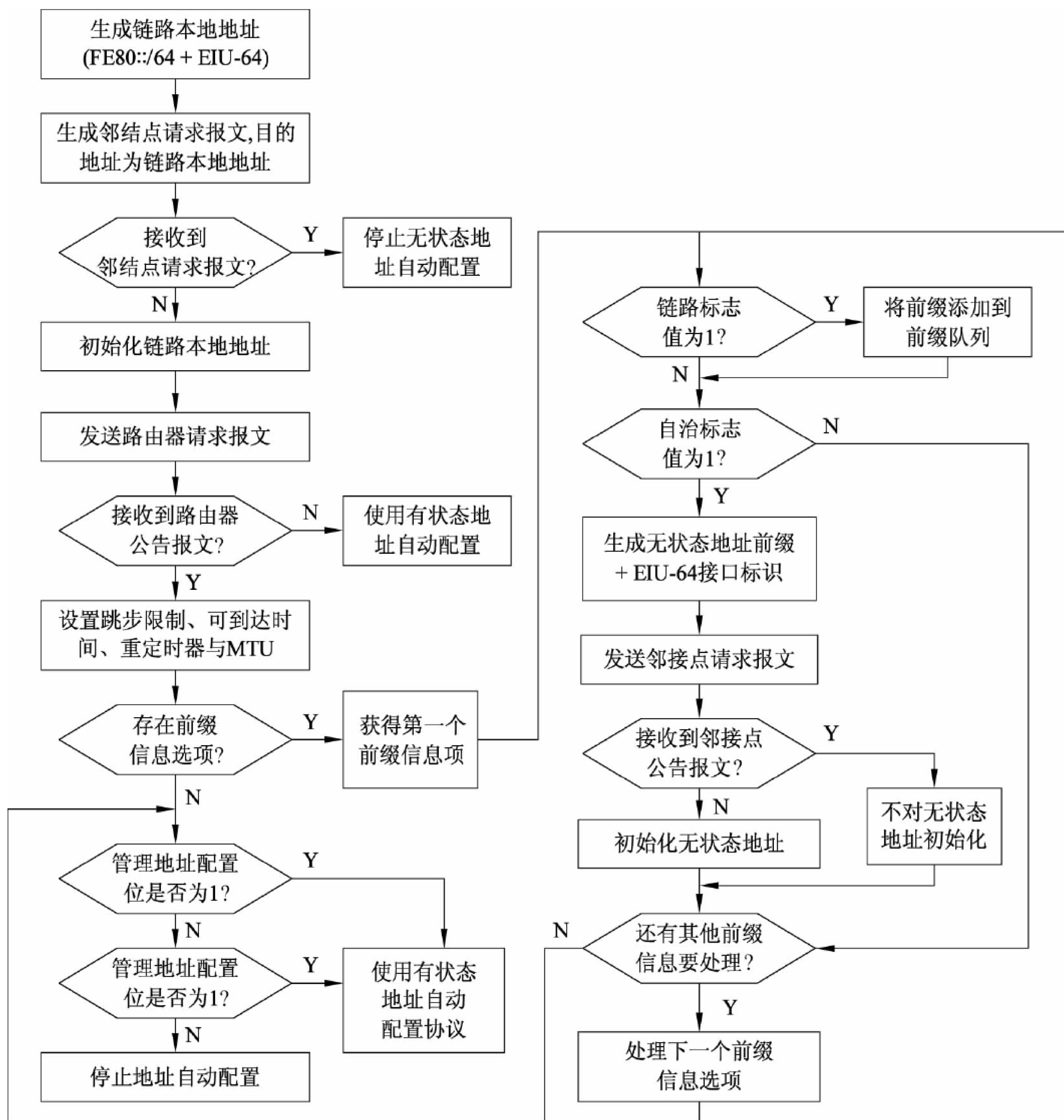


图5-35 IPv6无状态地址自动配置流程

3. IPv6有状态地址自动配置

有状态地址自动配置协议DHCPv6是一个基于UDP与Client/Server体系结构的协议,DHCPv6的工作模型如图5-36所示。

从上图中可以看出,执行有状态地址自动配置协议DHCPv6时,首先由主机发送一个DHCP Server Solicitation报文,给多播地址去发现DHCP Server,要求它发送应答报文。当DHCP Server接收到DHCP Server Solicitation报文之后,如果它允许主机使用地址与其他配置参数时,它将返回一个Unicast Reply报文。该报文将包含主机IPv6地址与配置

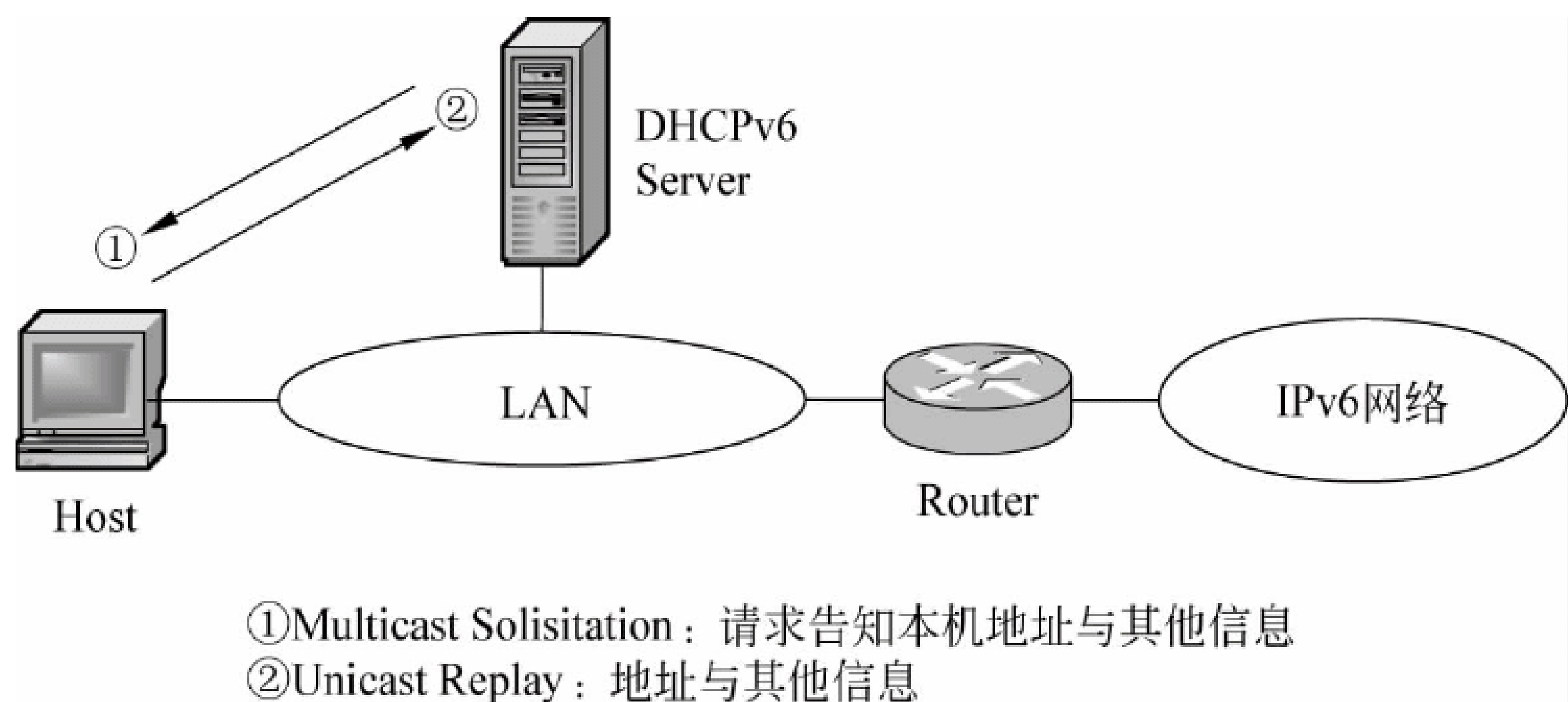


图 5-36　DHCPv6 的工作模型示意图

参数。

DHCPv6 可以使用两种多播地址：FF02∷1:2 用于同一链路内的主机向 DHCPv6 服务器发送 Multicast solicitation 报文；FF05∷1:2 用于同一站点内的主机与 DHCPv6 服务器通信。

问题 5-28：IP 地址习题有哪几种基本类型？

网络层是计算机网络课程学习中非常重要的内容之一，也是掌握计算机网络工作原理关键的内容，其中，IP 地址是这一章学习的重点，也是各种考核必考的内容。但是在学习的过程中，学生都感觉内容庞杂，很难理出头绪；考核时形式变化多样，灵活性太大，属于学习的难点，因此梳理一下这一部分内容的脉络是很有必要的。

总结在教学过程中以及布置作业时考虑的问题，大致可以将与 IP 地址相关的习题分为以下几种基本类型。

(1) 根据 IPv4 标准地址划分的方法，判断 A 类、B 类与 C 类地址。

(2) 根据特殊地址分配的规定，判断特殊 IP 地址。

(3) 根据给定的地址与掩码，判断哪些地址属于一个子网。

(4) 根据用户要求，将给定的地址段划分出合适的子网，确定子网地址位数、子网地址区间、子网广播地址、子网可用地址。

(5) 将给定的多个地址进行聚合。

(6) 根据规定的网络结构与网络地址，填写路由表。

(7) 根据路由表和目的 IP 地址，找出输出路由。

需要注意的是，我们只是根据多年给学生布置作业、出考题时考虑的几种基本问题的类型，而每种类型的问题只要改变一下地址等数据，学生就需要重新算一遍。只要将其中的问题联系起来，就会出现很多新的问题。IP 地址问题既是学习网络原理的重点，又是很容易命题的考点，因此老师在教学过程中必须重视这一部分内容的教学和训练。

问题 5-29：术语辨析：缆段、网络、子网与互联网络。

术语"缆段""子网""网络"与"互联网络"涉及 MAC 层与网络层的一些重要的概念，也是初学者容易混淆的术语。

(1) 缆段(Segment)术语来自局域网，尤其是早期 Ethernet 的 10BASE-2、10BASE-5 标准中，用粗同轴电缆或细同轴电缆组网时，一根连接了多台主机的同轴电缆叫作一个"缆

段”。同在一个“缆段”的主机共享一个冲突域。

(2) 网络(Network)是一个含义最不确定,但是又最常用的术语。网络可以是覆盖一个国家的广域网,也可以是覆盖一个城市范围的城域网,覆盖一个教学楼、一个实验室的局域网,也可以是我们身边10m以内的个人区域网。网络既可以是单一类型的广域网、城域网或局域网,也可能是几种网络互联的网络系统。因此,理解“网络”术语的含义需要依据它所描述的环境和对象而定。

(3) 子网(Subnet)是单词“Subnetwork”的缩写。子网是网络的一部分,也可以理解为互联网络的一部分。但是,术语“子网”更多是出现在IP地址划分的讨论中。对应术语“子网”的是“子网号(Subnet-ID)”。一个子网是由路由器互联形成的网络系统的一个基本组成单元。一个子网中的所有主机的网络号与子网号必须是相同的。

(4) 如果将术语“互联网络(Internet或Internetwork)”与“互联网(Internet)”放在一起讨论时,就会发现,“互联网(Internet)”是专用名词,而“互联网络(Internet或Internetwork)”不是专用名词。“互联网络”是由路由器将多个网络互联起来的统称。

问题5-30:路由选择算法与路由选择协议是同一件事吗?

初次接触网络层路由问题时,经常认为路由选择算法与路由选择协议是一件事,只是两种不同的叫法。实际上,路由选择算法与路由选择协议的概念是不同的。路由选择算法的目标是产生一个路由表,为路由器转发IP分组找出适当的下一跳路由器,而设计路由选择协议的目标是实现路由表路由信息的动态更新。

问题5-31:路由选择算法的研究经历了怎样的发展过程?

这是很多初学者很自然提出的问题。要回答这个问题,需要回顾Internet走过的一段历史,然后再进入路由选择算法的讨论。

(1) 实际上,只要看过IPv4最初文档的研究人员就会发现,IPv4协议对路由选择没有做任何规定和限制,它只提出了直接交付、间接交付与路由选择的术语,其他问题并没有提出解决方法。正是因为这样,20世纪70、80年代有大量的学者在研究路由选择算法问题,也发表了很多文章,出版过多部专著与综述性的著作。读者在20世纪70、80年代跟踪计算机网络与Internet技术发展过程中,看过很多这方面的文章,这与当今大家都在研究Ad Hoc与WSN路由问题很相似。大量研究工作展开之后,人们最终发现,用复杂的方法去处理复杂的问题是没有出路的,只有用简单的方法去处理复杂问题,才有可能找到令我们基本满意的答案。路由选择算法的研究发展过程正说明这样一个朴素的道理。这就引起了之后在Internet中应用的分层路由与自治系统的解决思路。

(2) 一个理想的路由选择算法应具有的特点。

路由选择的核心是路由选择算法,路由选择算法为完善路由表提供了算法依据。一个理想的路由选择算法应具有如下一些特点。

① 算法必须是正确、稳定和公平的。

沿着路由表所指引的路径,分组一定能够最终到达的目的网络和目的主机。在网络通信量和网络拓扑相对稳定的情况下,路由算法应收敛于一个可以接受的解。算法应对所有用户都是平等的。

② 算法应该尽量简单。

路由选择算法的计算必然要耗费路由器的资源，增加分组转发的延时，算法只有尽量简单，才可能有实用价值。

③ 算法必须能够适应网络拓扑和通信量的变化。

网络拓扑与网络通信量的变化是必然的。当某个路由器或通信线路发生故障时，算法应能及时地改变路由。当网络的通信量发生变化时，算法应能自动改变路由，以均衡各链路的负载。这种自适应性表现出路由选择算法的“稳健性”。

④ 算法应该是最佳的。

算法的“最佳”是指以低的开销转发分组。衡量开销的因素可以是链路长度、数据速率、链路容量、保密、传播延时与费用等。正是因为需要考虑很多因素，因此不存在一种绝对的最佳路由算法。“最佳”是指相对于某一种特定条件和要求，给出的较为合理的路由选择。

(3) 路由选择算法考虑的主要参数。

在讨论路由选择算法时，将会涉及以下几个参数。

① 跳数(Hop Count)。

跳数是指一个分组从源结点到达目的结点经过的路由器的个数。一般来说，跳数越少的路径越好。

② 带宽(Bandwidth)。

带宽指链路的传输速率，例如，T1 链路传输速率为 1.544Mbps，也可以说 T1 链路的带宽为 1.544Mbps。

③ 延时(Delay)。

延时是指一个分组从源结点到达目的结点花费的时间。

④ 负载(Load)。

负载是指通过路由器或线路的单位时间通信量。

⑤ 可靠性(Reliability)。

可靠性是指传输过程中的误码率。

⑥ 开销(Overhead)。

开销一般是指传输过程中的耗费，耗费通常与所使用的链路带宽相关。

对于一个实际的路由选择算法，应尽可能接近于理想的算法。在不同的应用条件下，可以有不同的侧重。应当指出，路由选择是个非常复杂的问题，因为它涉及网络中的所有主机、路由器、通信线路。同时，网络拓扑与网络通信量随时在变化，这种变化事先无法知道。当网络发生拥塞时，路由选择算法应具有一定的缓解能力，但恰好在这种条件下，很难从网络中的各结点获得所需的路由选择信息。由于路由选择算法与拥塞控制算法直接相关，因此只能寻找出对于某种条件相对合理的路由选择。

问题 5-32：为什么要采取自治系统与分层路由的方法？

回答这个问题，需要注意以下几点。

(1) 对于结构极为复杂，并且结构不断变化的 Internet 来说，要想出现一个集中式、能够指导全网路由的路由选择算法是不可能的，因此必须采取“化整为零，分而治之”的思路来处理。

(2) Internet 采用分层的路由选择协议，并且将整个 Internet 划分为许多较小的自治系

统(Autonomous System,AS)。一个自治系统内的所有网络都属于一个行政单位,例如一所大学、一个公司、政府的一个部门。

(3) 一个自治系统最重要的特点就是它有权自主地决定在本系统内应采用何种路由选择协议。作为一个自治系统,其核心是路由寻址的"自治"。它应该包括以下两个方面的内容。

① 自治系统内部的路由器了解内部全部网络的路由信息,并能通过一条路径将发送到其他自治系统的分组传送到连接本自治系统的主干路由器。

② 自治系统内部的路由器要向主干路由器报告内部路由信息。

(4) Internet将路由选择协议分为两大类:内部网关协议(Interior Gateway Protocol,IGP)与外部网关协议(Exterior Gateway Protocol,EGP)。

了解分层路由与自治系统的概念,对深入理解Internet网络层实现技术是十分重要的。

问题5-33:如何认识路由选择协议的特点?

Internet将路由选择协议分为两大类。

(1) 内部网关协议是在一个自治系统内部使用的路由选择协议,这与Internet中的其他自治系统选用什么路由选择协议无关。目前,内部网关协议主要有:路由信息协议(Routing Information Protocol,RIP)和开放最短路径优先(Open Shortest Path First,OSPF)协议。

当源主机和目的主机处在不同的自治系统中,并且这两个自治系统使用不同的内部网关协议时,那么当分组传送到一个自治系统的边界时,就需要使用一种协议将路由选择信息传递到另一个自治系统中,这时需要使用外部网关协议。目前,外部网关协议主要是边界网关协议(Border Gateway Protocol,BGP)。

(2) 每个自治系统运行用于内部路由选择的IGP,但每个自治系统都有一个或多个与其他自治系统连接的路由器,除了运行本自治系统的IGP以外,还需要运行用于自治系统之间路由选择的EGP。在各个自治系统内使用的IGP可以是RIP,也可以是OSPF。在自治系统之间必须使用EGP(例如BGP)。

(3) 在研究Internet路由选择协议时,需要注意以下几个问题。

① 路由选择算法和路由选择协议在概念上是不同的。网络上的主机、路由器通过路由选择算法形成路由表,以确定发送分组的传输路径。而路由选择协议是路由器用来完成路由表建立和路由信息更新的通信协议。

② 早期RFC文档中使用的术语"网关(Gateway)",相当于今天人们非常熟悉的"路由器(Router)"。新的RFC文档中使用了"路由器"。从网络互联设备角度来看,"网关"与"路由器"是有区别的,但是由于历史的原因,在Internet的路由技术的讨论中,"网关"与"路由器"没有加以区别。

③ IGP与EGP是两类Internet路由选择协议的名称,但是早期一种具体的关于外部网关协议就叫作"EGP"(RFC827)。因此,出现了一类协议的名称与其中一种具体的协议名称相同的情况,容易造成混淆。此后,出现了一种新的外部网关协议——边界网关协议(RFC1771、RFC1772),它取代了RFC827的EGP,成为目前广泛使用的一种具体的边界网关协议。

问题 5-34：RIP 与 OSPF 协议的区别是什么？

RIP 与 OSPF 协议的区别主要表现在以下几个方面。

(1) OSPF 使用的是链路状态协议(Link State Protocol)，而 RIP 使用的是向量-距离路由选择协议。

(2) OSPF 要求每个路由器周期性发送链路状态信息，以使区域内的所有路由器最终都能形成一个跟踪网络链路状态的链路状态数据库(Link State Database)，这些状态包括路由器可用端口、已知可达路由和链路状态信息。实际上，链路状态数据库是一张完整的网络映射图，它是路由器建立路由表的依据。RIP 只能根据相邻路由器的信息更新路由表。

(3) OSPF 要求路由器在链路状态发生变化时用洪泛法(Flooding)向所有路由器发送该信息，而 RIP 仅向自己相邻的几个路由器通报路由信息。

(4) 路由器之间交换的链路状态信息主要指费用、距离、延时、带宽等。注意，RIP 与 OSPF 都是在寻找最短的路径，并且都采取“最短路径优先”的指导思想。但是，在具体使用怎样的参数，以及计算方法上有些不同。RIP 采用的是“跳数”作为路径长短的度量，它是以跳数最少作为“最短路径”的评价标准。

问题 5-35：为什么不对“头部校验和出错”发送 ICMP 差错报告报文？

在路由器接收到一个 IP 分组后，检查分组头校验和发现出错时，并不向丢弃分组的源主机发送 ICMP 差错报告报文，其理由有两点。一是 IP 协议不要求源主机在出现传输错误时必须重传，而发现数据报文是否出错以及出错后如何处理将由高层协议解决；二是分组头校验和发现出错不能保证不是源 IP 地址出错。如果是源 IP 地址出错，那么向可能出错的源主机发送 ICMP 差错报文也是没有意义的。

问题 5-36：如何理解 IP 多播的基本概念？

理解 IP 多播的基本概念，需要注意以下几个问题。

1. IP 多播发展的过程

传统的 IP 协议规定 IP 分组的目的地址只能是一个单播(Unicast)地址。这种 IP 单播工作模式对于新闻、股市与金融信息发布，以及讨论组、视频会议、交互式游戏等由多个用户参与的网络应用，显然会大量浪费网络资源，并且工作效率很低。

1988 年，Steve Deering 首次提出 IP 多播(或组播)的概念。1989 年，RFC1112 对 IP 多播协议(IGMP)进行定义。为了适应交互式音频和视频信息的多播，从 1992 年起开始实验虚拟的多播主干网(Mbone)。Mbone 可以将分组发送到属于一个组的多个主机。RFC2236 文档是 1997 年公布 IGMP 的第二个版本，已成为互联网标准协议。RFC3376 对 IGMP 组管理协议进行讨论。目前多播主干网的规模已经很大，拥有几千个多播路由器。在 IP 协议的框架之下，传输多播分组也一定只能提供“尽力而为”服务。IGMP 不能保证多播分组一定能到达所有的组成员。

2. IP 多播与单播的区别

在 IP 单播状态下，如果主机 0 打算向主机 1～主机 20 发送同一文件，则它需要准备 20 个文件的副本，分别封装在源地址相同，而目的地址不同的 20 个分组中，分别将这 20 个分组发送给 20 个目的主机。

在 IP 多播状态下，如果主机 0 打算向多播组成员主机 1～主机 20 发送同一文件，则它

只需要准备一个文件的副本,封装在一个多播分组中,发送给多播组中20个多播组成员。如果IP多播组的成员达到成千上万个时,多播工作对系统效率的提高将会更加显著。支持IGMP的路由器称为多播路由器(Multicast Router)。

问题5-37:IP多播地址是如何规定的?

在讨论IP多播地址时,需要注意以下几个问题。

(1) 实现IP多播的分组使用的是IP多播地址。IP多播地址只能用于目的地址,而不能用于源地址。

(2) 标准分类的D类地址是为IP多播地址而定义的。D类IP地址的前4位为1110,因此D类地址的范围在224.0.0.0~239.255.255.255。每个D类IP地址可以用于标识一个多播组,则D类地址能标识出2^{28}个多播组。

(3) RFC3330说明D类地址空间中被预留作特殊用途的部分多播地址。例如,224.0.0.0被保留;224.0.0.1指定为本网中所有参加多播的主机和路由器使用;224.0.0.11指定为移动代理的地址;224.0.1.1~224.0.1.18被预留为电视会议等多播的应用;239.0.0.0~239.255.255.255限制在一个组织中使用。完整的保留多播地址表可以从IANA网站获取。

问题5-38:IGMP包括哪些基本内容?

理解这个问题需要注意以下几点。

(1) IP多播的基本思想是:多个接收者可以接收到从同一个或一组源结点一次发送的相同内容的分组。

(2) 发送IP多播工作模式包括以下内容。

① 定义了一个组地址(Group Address)。每个组代表一个或多个发送者与一个或多个接收者的一个会话(Session)。

② 接收者可以用多播地址通知路由器,它希望加入(或退出)哪个多播组。

③ 发送者使用多播地址发送分组,无须了解接收者的位置信息与状态信息。

④ 路由器建立一棵从发送者分支出去的多播传递树,这棵树延伸到所有的、其中至少有一个IP多播成员的网络中。利用这棵传递树,路由器把多播组的分组一直转发到有多播组成员的网络中。

(3) 当某个主机加入新的多播组时,该主机应向多播组的多播地址发送一个IGMP报文,声明自己要成为该组的成员。本地的多播路由器收到IGMP报文后,将组成员关系转发给互联网上的其他多播路由器。

(4) 由于多播组的成员关系是动态的,因此本地多播路由器要周期性地探询本网络上的主机,以便知道这些主机是否还继续是组的成员。只要对某个组有一个主机响应,则多播路由器就认为这个组是活跃的。如果一个组在经过几次探询后没有一个主机响应,则多播路由器就认为本网络上的主机都已离开这个组,不再将该组的成员关系转发给其他多播路由器。

多播路由器在探询组成员关系时,只需对所有的组发送一个请求信息的询问报文,而不需要对每个组发送一个询问报文。同一组内的每个主机都要监听响应,只要有本组的其他主机先发送响应,自己就可以不再发送响应。这样就抑制了不必要的通信量。多播路由器只需知道网络上是否至少还有一个主机是本组成员。

问题5-39：什么时候需要使用IP多播隧道技术?

当多播IP分组跨越多个网络时，存在关于多播IP分组的路由问题。多播路由器的作用是完成多播分组的转发工作，具体有两种实现方式：一种是专用多播路由器；另一种是在传统路由器上实现多播路由的功能。

在多播传送中，当多播路由器对多播分组进行存储转发时，在任一多播路由器所在的网络上都可能有该多播组成员，在传送过程中随时会遇到某个目的主机。这也是多播传送的一大特点。当IP多播分组在传输的过程中遇到有不支持多播协议的路由器或网络时，就要采用隧道技术。

问题5-40：为什么要研究RSVP、DiffServ与MPLS技术?

理解这个问题，需要注意以下几点。

(1) 网络中不同的层次都会涉及服务质量(QoS)问题。评价网络层QoS的参数主要是带宽与传输延时。IP协议提供的"尽力而为"服务，对于多媒体网络服务显然不适应。

(2) 在网络层引入QoS保障机制的目的是：通过协商为某种网络服务提供所需的网络资源，防止个别网络应用独占共享的网络资源。因此，QoS保障机制实际上是一种网络资源分配机制。

在讨论IP网络的QoS问题的背景之下，出现了资源预留协议、区分服务与多协议标记服务技术的研究。

问题5-41：什么是资源预留协议RSVP?

回答这个问题需要注意以下几点。

(1) 资源预留协议(RSVP)的核心是对一个应用会话的数据流提供服务质量保证。流的定义是"具有相同源IP地址、源端口号、目的IP地址、目的端口号、协议标识符与服务质量要求的分组序列"。

(2) 资源预留意味着路由器知道需要为即将出现的会话预留多少链路带宽和缓冲区。为了做到这一点，需要源结点和目的结点之间在会话之前建立一个连接，路径上的所有路由器都要预留需要的资源。

(3) 由于RSVP基于单个数据流的端-端资源预留、调度处理和缓冲区管理、状态维护机制太复杂，开销太大，不适用于大型网络。在目前的网络上推行RSVP服务，需要对现有的路由器、主机与应用程序做相应的调整，实现难度也很大。

(4) 单纯的RSVP结构实际上无法让业界接受，也无法在互联网上得到广泛的应用。

问题5-42：什么是区分服务DiffServ?

回答这个问题需要注意以下几点。

(1) RSVP应用的受阻，促进区分服务(DiffServ)技术的研究与发展。针对RSVP存在的问题，DiffServ的设计者注意解决协议的简单、有效与可扩展性问题，使它成为适用于骨干网的多种业务服务需求。

(2) DiffServ与RSVP的区别主要表现在以下几点。

① RSVP是基于某一个会话流，而DiffServ是基于某一类应用的。以IP电话为例，RSVP只为一对通话的用户提取建立一个连接，预约带宽与缓冲区，以保证这一对用户的通话质量。而DiffServ是针对IP电话这一类应用。如果ISP为IP电话设置为保证服务质量

的一类服务,那么IP电话的数据分组的服务类型字段就带有标记。IP电话的数据分组进入ISP网络时,网络就要为IP电话的数据分组提供高质量的传输服务。

② RSVP要求所有的路由器都要修改软件,以支持基于流的传输服务服务。DiffServ只需要一组路由器(如ISP网络中的路由器)支持,就可以实现DiffServ服务。

(3) 当IETF完成了RSVP与DiffServ协议的研究,有些路由器厂商又提出了更好的改善IP分组传输质量的方案,那就是多协议标识交换MPLS技术。

问题5-43:什么是多协议标识交换MPLS?

回答这个问题需要注意以下几点。

1. MPLS的基本概念

从设计思想上来看,MPLS将数据链路层的第二层交换技术引入网络层,实现快速IP分组交换。在这种网络结构中,核心网络是MPLS域,构成它的路由器是标记交换路由器(Label Switching Router,LSR),在MPLS域边缘连接其他子网的路由器是边界标记交换路由器E-LSR。MPLS在E-LSR之间建立标记交换路径(Lable Switching Path,LSP),这种标记交换路径LSP与ATM虚电路VC非常相似。MPLS减少IP网络中每个路由器逐个分组处理的工作量,可以进一步提高路由器性能和传输网络的服务质量。

2. MPLS可以提供的4个主要的服务功能

1) 提供面向连接与保证QoS的服务

MPLS的设计思路是借鉴ATM面向连接和可以提供QoS保障的设计思想,在IP网络中提供一种面向连接的服务。

2) 合理利用网络资源

流量工程(Traffic Engineering,TE)研究的目的是更合理地利用网络资源,提高服务质量。流量工程不是特定于MPLS的产物,而是一种通用的概念和方法,是拥塞控制研究中的均衡负荷方法。基于MPLS的流量工程是利用面向连接的流量工程技术与IP路由技术相结合,动态地定义路由。MPLS引入"流"的概念。流是从某个源结点发出的分组序列,利用MPLS可以为单个流建立路由。为满足不同流的服务质量需求,对端之间可以为不同的流选择不同的路由。

3) 支持虚拟专网服务

MPLS提供VPN服务,提高分组传输的安全性与服务质量。

4) 支持多协议

支持MPLS协议的路由器可以与普通IP路由器、ATM交换机、支持MPLS的帧中继(RF)交换机共存。因此,MPLS可以用于纯IP网络、ATM网络、帧中继网络及多种混合型网络,同时可以支持PPP、SDH、DWDM等多种底层网络协议。

问题5-44:MPLS VPN具有哪些特点?

理解这个问题需要注意以下几点。

1. VPN的实现方法

(1) 通过在帧中继网、ATM网或IP网上配置一条或多条加密的点-点隧道的方法为用户建立VPN。网络管理人员需要根据不同的用户需求,设置每一条隧道或虚电路的参数。在传统的VPN技术中,很多需要构建企业专网的公司通过租用专用线路或帧中继链路来

建立第二层 VPN(L2VPN),以满足企业对数据传输安全的需要。

(2) 在 MPLS 网络中提供 VPN 服务。MPLS 可以将面向连接的标记路由机制与 VPN 的建设需求结合起来,为所有连入 MPLS 网络的用户之间方便地建立第三层 VPN (L3VPN)。

2. MPLS VPN 的特点

在基于 MPLS 的 VPN 中,服务提供商为每个 VPN 分配一个路由标识符(RD)。这个路由标识符在 MPLS 网络中是唯一的。标记交换路由器 LSR 和边界路由器 E-LSR 的标记转发表中记录了该 VPN 中用户 IP 与路由标识符 RD 的对应关系。只有属于同一个 VPN 的用户之间才能通信。保证数据传输安全性。显然,MPLS VPN 技术可以满足用户关于保证数据通信安全性、网络服务质量、操作的简便性与可扩展性的要求。因此,在大型信息网络系统、物联网应用系统、云计算系统中 MPLS VPN 技术已经得到广泛应用。

问题 5-45:为什么要研究 ARP 技术?

回答这个问题需要注意以下几点。

(1) 对于 TCP/IP 来说,主机和路由器在网络层用 IP 地址来标识,在数据链路层用物理地址(例如在 Ethernet 的 MAC 地址)来标识。在描述一个网络的工作过程时,实际上是做了一个假设:已经知道通信的目的主机的 IP 地址,并且知道对应这个 IP 地址的目的主机物理地址。这个假设成立的条件是:在任何一台主机或路由器中必须有一张“IP 地址-物理地址对照表”,它应该包括所需通信的任何一台主机或路由器的信息。

(2) 通过“静态映射”的方法,从一个已知的 IP 地址获取与之对应的物理地址。但是,这是非常理想的一种解决方案,在一个小型的互联网络系统中实现起来比较容易,这在大型网络中几乎是不可能实现的。这种方法有很大的局限性。

① 如果有一个主机或路由器刚刚加入到网络中,其他结点的“IP 地址-物理地址对照表”不会有它的信息。

② 如果一个主机更换了网卡,在 IP 地址不变的情况下,它的物理地址发生了改变。

③ 不同物理网络的网络地址结构、长度与设置方法都可能是不一样的。有些局域网(例如 LocalTalk)每次加电时,其物理地址都要改变一次,那么这种网络的 IP 地址与物理地址没有确定的对照关系。

④ 如果一个主机从一个物理网络移到另一个物理网络,那么它的物理地址不变,而 IP 地址发生了变化。

因此,在 Internet 中应该设计一种“动态映射”的方法来解决 IP 地址与物理地址映射的问题。

(3) 从已知的 IP 地址找出对应的物理地址的映射过程称为正向地址解析,相应的协议称为地址解析协议(ARP)。从已知的物理地址找出对应的 IP 地址的映射过程称为反向地址解析,相应的协议叫作反向地址解析协议(RARP)。

图 5-37 给出了 ARP 在网络实现技术中的作用示意图。

问题 5-46:ARP 功能是如何实现的?

要了解 ARP 功能的实现方法,需要了解 ARP 请求与应答分组。

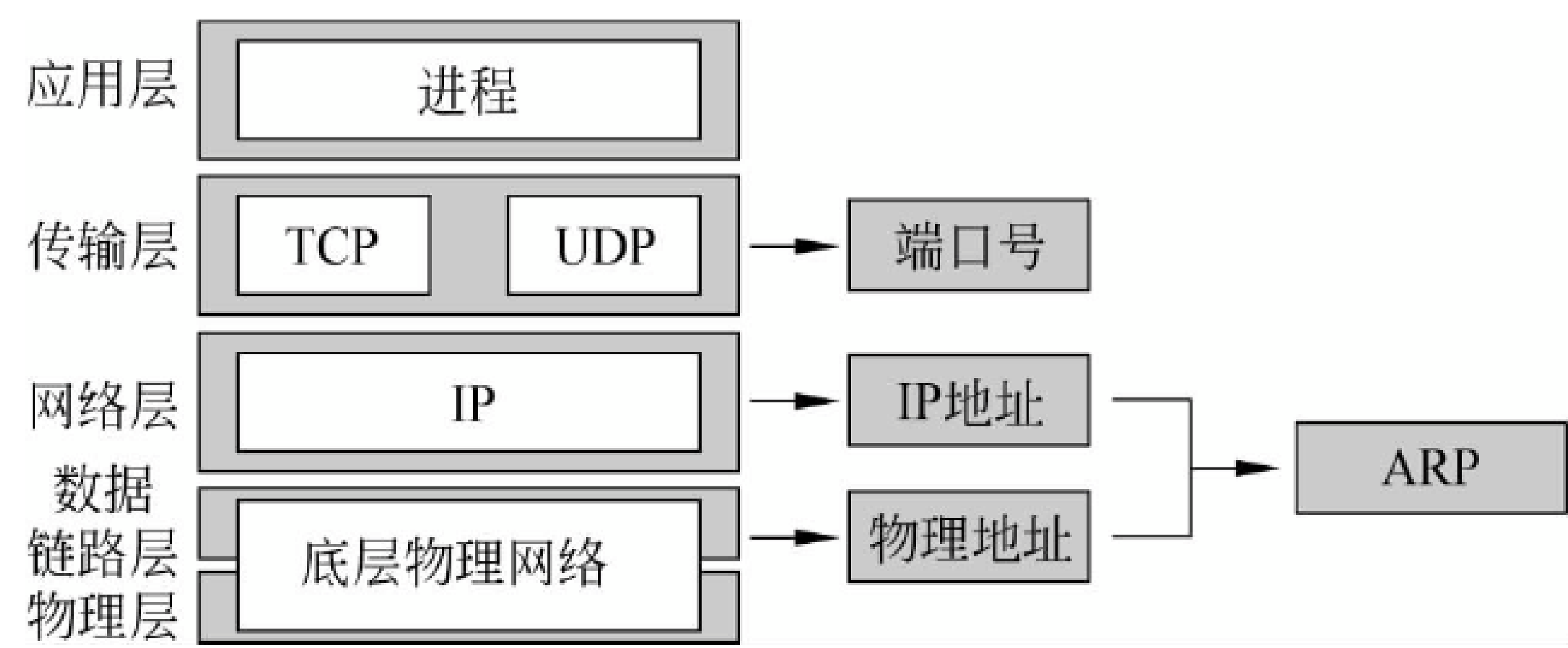


图 5-37　ARP 在网络实现技术中的作用

1. ARP 请求与应答分组格式

在地址解析协议中,ARP 请求分组与应答分组的格式如图 5-38 所示。

0　　　　8	16	24　　　　31
硬件类型		协议类型
硬件地址长度	协议长度	操作
发送结点硬件地址		
发送结点硬件地址		发送结点协议地址长度
目的结点硬件地址		
目的结点硬件地址		目的结点协议地址长度

图 5-38　ARP 分组格式

2. ARP 分组中各个字段的作用

1) 硬件类型

硬件类型字段长度为 16b,表示发送端物理网络的类型。例如,硬件类型字段数值为 1 时,表示发送端协议是 Ethernet。ARP 分组允许使用各种网络,它具体地为每一种类型的局域网分配了一个代码。

2) 协议类型

协议类型字段长度为 16b,表示发送端网络层的协议类型。例如,该字段值为 0x0800,则表示发送端采用的是 IPv4 协议。

3) 硬件长度

硬件长度字段长度为 8b,表示以字节为单位的物理地址的长度。例如,硬件长度值为 6 是 Ethernet 地址。

4) 协议长度

协议长度字段长度为 8b,表示以字节为单位的网络层地址的长度。例如,协议长度字段的值为 4 表示 IPv4 协议。

5) 操作

操作字段长度为 16b,表示分组的类型。如果该字段数值为 1,则表示 ARP 请求分组;如果该字段数值为 2,则表示 ARP 响应分组;如果该字段数值为 3,则表示 RARP 请求分组;如果该字段数值为 4,则表示 RARP 响应分组。

6) 发送端硬件地址

发送端硬件地址字段长度可变,表示以字节为单位的源结点的物理地址长度类型。

Ethernet 的硬件长度字段的值为 6。

7）发送端协议长度

发送端协议长度字段长度为 8b，表示以字节为单位的源结点网络层地址的长度。例如，发送端协议长度值为 4 是 IPv4 协议。

8）目的端硬件地址

目的端硬件地址字段长度可变，表示以字节为单位的发送站的物理地址长度类型。Ethernet 的硬件长度字段的值为 6。对于 ARP 请求分组，由于不知道目的物理地址长度，因此用全 0 表示。

9）目的端协议长度

目的端协议长度字段长度为 8b，表示用字节为单位的网络层地址的长度。目的端协议长度值为 4，表示为 IPv4 协议。

问题 5-47：ARP 基本工作过程是怎样的？

地址解析一般是将静态映射与动态映射的方法结合起来。人们首先要在本地主机内部建立一个“ARP 高速缓存表”，用来存储部分 IP 地址与物理地址映射关系。该表可以随着时间而动态更新。

1. ARP 执行的过程

在考虑到地址解析过程后，一个结点在发送分组时需要经过以下几个步骤。

(1) 如果主机 A 打算给主机 B 发送一个 IP 分组，它知道主机 B 的 IP 地址，但是不知道主机 B 的 MAC 地址。那么它首先要在本地 ARP 映射表中查找。如果找到，就不需要进行地址解析。如果查找不到，则需要进行地址解析。

(2) 地址解析第一步是由主机 A 产生 ARP 请求分组，在相应的字段写入主机 A 的源 MAC 地址与源 IP 地址，主机 B 的 IP 地址作为目的 IP 地址，在目的 MAC 地址字段写入 0。

(3) 将 ARP 分组传递到本地的数据链路层，并组装成 ARP 请求分组的帧。源 MAC 地址是发出 ARP 请求分组的主机 MAC 地址，目的地址是广播地址(ff-ff-ff-ff-ff-ff)，通过物理层发送出去。

(4) 封装了 ARP 请求分组的帧通过广播方式发送出去，包括主机 B 在内的所有的主机都能接收到 ARP 请求分组。接收到 ARP 请求分组的主机，如果它的映射表中没有主机 A 的 IP 地址对应的 MAC 地址，那么它就将主机 A 的 IP 地址、MAC 地址对应关系存入映射表。每台主机都可以通过接收到 ARP 请求分组来不断完善它的映射表。

(5) 主机 B 在接收到主机 A 的 ARP 应答分组之后，就向主机 A 发送一个封装了 ARP 应答分组的帧。ARP 应答分组包含主机 B 的 IP 地址、MAC 地址。

(6) 主机 A 在收到 ARP 应答分组之后，将主机 B 的 IP 地址、MAC 地址存入到映射表。这样，主机 A 获得了主机 B 的 IP 地址对应的 MAC 地址，它就可以直接向主机 B 发送数据帧。

2. ARP 缓存

ARP 高效运行的关键是在每个主机和路由器上维护一个 ARP 缓存表。该缓存使用地址解析为每个接口维护从网络层地址到硬件地址的最新映射。RFC1122 规定：当 IPv4 地址映射到硬件地址时，高速缓存中一个条目的正常到期时间是条目创建开始后 20 分钟。

读者可在 Linux 或 Windows 中使用 arp 命令查看 ARP 缓存。选项-a 用于显示任何系

统的缓存中的所有条目。在 Linux 中,运行 arp 命令会产生以下输出。

```
Linux%arp
Address         HWtype  HWaddress           Flags     Mask Iface
gw.home         ether   00:0D:66:4F:60:00   C         eth1
printer.home    ether   00:0A:95:87:38:6A   C         eth1
Linux%arp -a
printer.home(10.0.0.4)at 00:0A:95:87:38:6A [ether] on eth1
gw.home(10.0.0.1)at 00:0D:66:4F:60:00 [ether] on eth1
```

在 Windows 中,运行 arp 会产生以下类似的输出:

```
c:\>arp -a
Interface: 10.0.0.56 --- 0x2
Internet Address    Physical Address     Type
10.0.0.1            00-0d-66-4f-60-00    dynamic
10.0.0.4            00-0a-95-87-38-6a    dynamic
```

这里可以看到的是 IPv4 到硬件地址的缓存。在第一种(Linux)情况下,每个映射是一个包含 5 个元素的条目:主机名(对应一个 IP 地址)、硬件地址类型、硬件地址、标志和本地网络接口,它对于这个映射是活跃的。标志列包含一个符号:C、M 或 P。C 类条目由 ARP 动态学习。M 类条目通过手工输入(arp -s)。P 类条目的含义是“发布”。也就是说,对于任何 P 类条目,主机对输入的 ARP 请求返回一个 ARP 应答。这个选项用于配置代理 ARP。第二个 Linux 的例子显示了使用“BSD 风格”的类似信息。这里,无论是主机名还是获得的地址,取决于地址类型(ether 表示一个以太网类型的地址),以及映射在哪个接口上是活跃的。

Windows 的 arp 程序显示了接口的 IPv4 地址,它的接口号是十六进制数(0x2)。Windows 版本指出地址是手动输入还是 ARP 学习。在这个例子中,两个条目都是动态的,这意味着它们来自 ARP 学习。通过手工输入,它们是静态的。注意,48 位 MAC 地址被显示为 6 个十六进制数,在 Linux 中使用冒号分隔,在 Windows 中使用破折号分隔。UNIX 系统一直使用冒号,而 IEEE 标准和其他操作系统使用破折号。

问题 5-48:IP 分组在转发过程中 IP 地址与 MAC 地址到底哪个在变?

理解 ARP,对于理解 Internet 工作原理是非常有帮助的。学生在讨论 IP 分组头时经常会提问:“IP 分组在转发过程中,到底是 IP 地址变,还是 MAC 地址在变?”如果将 ARP 的设计目的、工作原理与执行过程与这个问题结合起来讨论,我们会发现:搞清楚这个问题,对于理解 Internet 工作原理非常有帮助。

我们假设一种情况:如果通过一台计算机访问一台国外大学 Web 服务器,那么在通信之前,只可能知道它的服务器的域名,通过域名解析知道对应 IP 地址是可能的,但是要我们同时也能找到这台 Web 服务器域名对应的 MAC 地址是不合理的。这也不符合 IP 协议设计的基本原则。按照 IP 协议的工作原理,它需要根据源 IP 地址与目的 IP 地址,寻找合适的传输路由,通过一跳一跳的路由器来接续,将 IP 分组转发到目的主机。同时,我们也不可能了解传输路径上所有路由器与目的 Web 服务器的 MAC 地址,这个转发过程协议是由 ARP 自动完成的。ARP 的执行过程对用户是透明的。

我们可以将 ARP 按照 IP 协议执行过程(直接交付-间接交付-直接交付)的思路画出来(如图 5-39 所示),那么就会发现:路由器每次转发 IP 分组时,分组头的源 IP 地址与目的 IP 地址是不变的,改变的是帧的源 MAC 地址与目的 MAC 地址。



图 5-39　分组在转发过程 IP 地址与 MAC 地址的变化

问题 5-49:如何认识中继器、集线器、网桥、交换机、路由器与网关的区别?

计算机网络由主机、网络设备与传输介质组成。网络设备可以是中继器(Repeater)、集线器(Hub)、网桥(Bridge)、交换机(Switch)、路由器(Router)或者是网关(Gateway)。它们分别工作在物理层、数据链路层、网络层,以及高层。网络设备与对应的工作层次关系如图 5-40 所示。

层次	网络设备
应用层	Gateway
传输层	
网络层	Router
数据链路层	Bridge
物理层	Repeater　Hub

图 5-40　网络设备与对应的层次关系示意图

实际上,在计算机网络研究与互联网形成、发展的过程中,不同的阶段对于功能与我们目前使用的路由器相同的网络连接设备的名称有多个,如 ARPANET 中出现的“接口报文处理机(IMP)”,有的文件中也将它称作“通信控制处理机(CCP)”。第一个讨论 IPv4 的

RFC791文档是1981年9月公布的,当时研究人员用的术语是“IP网关”,因此这个术语在早期计算机网络的讨论中经常出现。到20世纪90年代,网络设备类型越来越多,功能也越来越明晰,一些重要的网络产品制造商开始意识到有必要将路由器(Router)与网关(Gateway)区分开。Router作为实现IP协议中路由算法、分组交付的主要网络层设备,其功能非常清晰,而Gateway作为一种协议变换设备可以工作在传输层,也可以工作在应用层。因此,如果我们限定在网络层,讨论RFC791描述的IPv4协议时,“IP网关”与“IP路由器”的概念与功能应该是相同的。

问题5-50:路由器要接入Ethernet、FE、GE与PPP在硬件上应该如何处理?

这对于理解路由器在互联网中的作用与工作原理是非常关键的问题。图5-41给出了解释这个问题的路由器结构示意图。

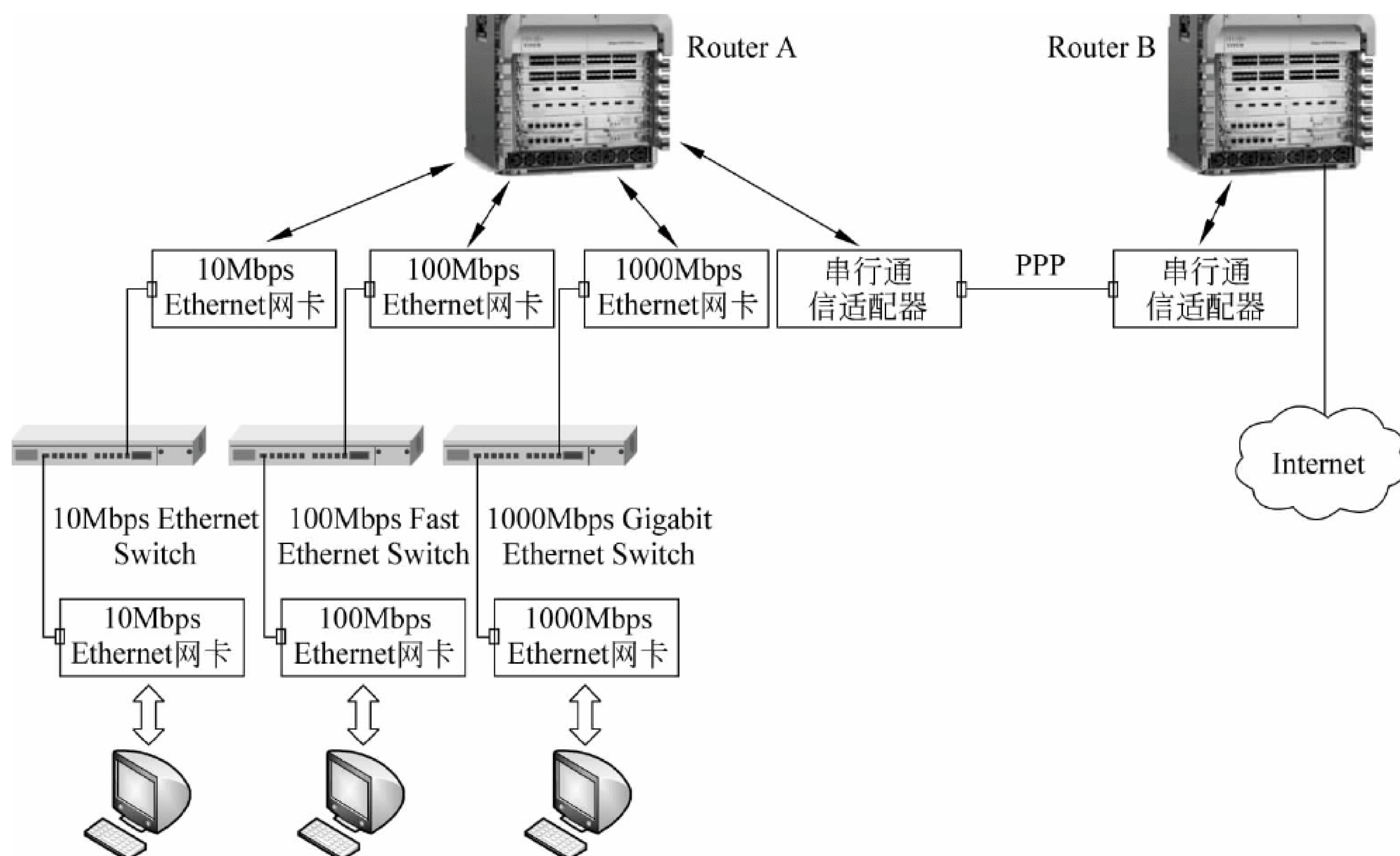


图5-41 路由器硬件结构示意图

在路由器互连两个系统的Ethernet的情况中,如果要求路由器同时能够接入10Mbps的Ethernet、100Mbps的FE与1Gbps的GE,实际上需要在路由器中增加支持不同协议的网卡,如图5-41中Router A有连接传统Ethernet以及FE、GE的网卡。实际组网过程中,需要根据所要接入的网络类型,在路由器的选型中选择支持不同网络协议(如Ethernet、FE、GE、10GbE、串行通信等)的模板。在图5-41中,选择了可以用于光纤连接的串行通信适配器模板,那么就可以使用PPP,通过光纤将两台路由器互连起来。这些知识对于理解网络设计与组网原理是很重要的。

问题5-51:路由器有哪几种基本类型?

路由器可以按照性能、部署的位置、使用的信道类型进行分类(如图5-42所示)。

1. 高端路由器与中低端路由器

按照性能可以将路由器分为高端路由器与中低端路由器。实际上,各个路由器生产商的分类标准是不同的。通常人们将背板交换能力大于40Gbps的路由器称为高端路由器,

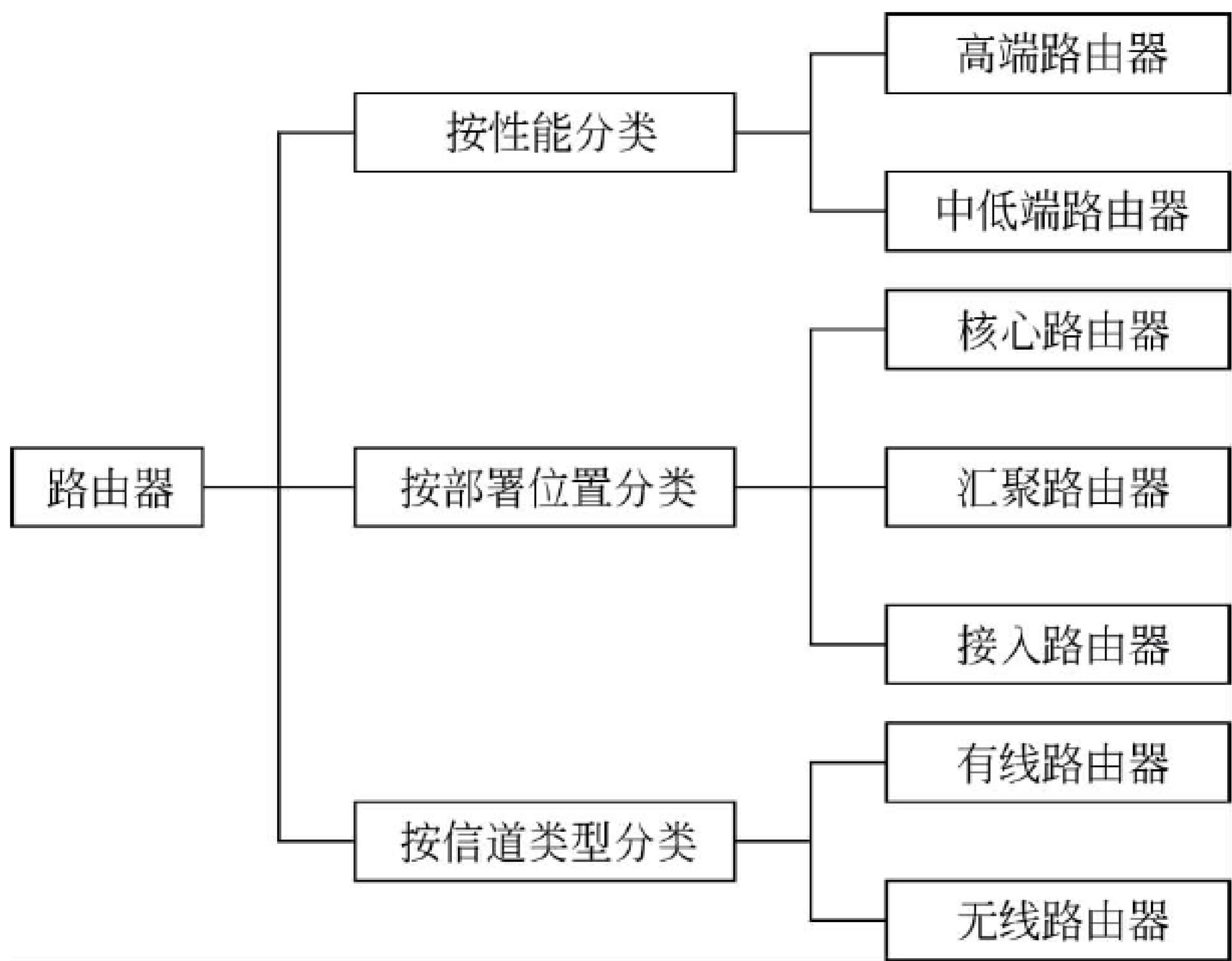


图 5-42　路由器的分类

背板交换能力低于 40Gbps 的路由器称为中低端路由器。

2. 核心路由器、汇聚路由器与接入路由器

按照部署的位置可以将路由器分为核心路由器、汇聚路由器与接入路由器。从传输网层次结构的角度，核心路由器是构成主干网的路由器，它仅需要支持 IP 协议。在实际应用中，人们又进一步地将核心路由器分为 Gbps 路由器、Tbps 路由器与交换路由器等。

汇聚路由器用于将大量的接入路由器汇聚到主干网。核心路由器只需要支持 IP 协议，而汇聚路由器要面对不同的接入网络，因此它需要考虑对不同协议类型的支持，以及对防火器、流量均衡设备、网络安全设备、MPLS、VPN 功能的支持。

一般来说，核心路由器要选用高端路由器。但是，随着网络规模的不同，以及路由器性能的不断提高，目前的核心路由器有可能很快成为汇聚路由器。路由器的分类是相对的，它是在不断变化的。

3. 有线路由器与无线路由器

按照使用的信道类型可以将路由器分为有线路由器与无线路由器。由于移动互联网的大规模应用，越来越多的办公室、实验室、家庭都在使用无线路由器。相对于无线路由器，目前互联网大量使用的路由器都应该属于有线路由器，只是我们已经习惯将它简称为路由器而已。无线路由器一般都支持专线 ADSL、Ethernet 等多种接入方式，还可以提供 DHCP、NAT、防火墙与 MAC 地址过滤等服务。

问题 5-52：如何认识路由器的结构与工作原理？

1. 路由器的基本结构

路由器是一种具有多个输入/输出端口，完成分组转发功能的专用计算机系统。它是由“路由选择处理机”和“分组处理与交换”两部分组成的。图 5-43 给出了典型的路由器结构示意图。

1）路由选择处理机

路由选择处理器是路由的控制部分，它的任务是生成和维护路由表。

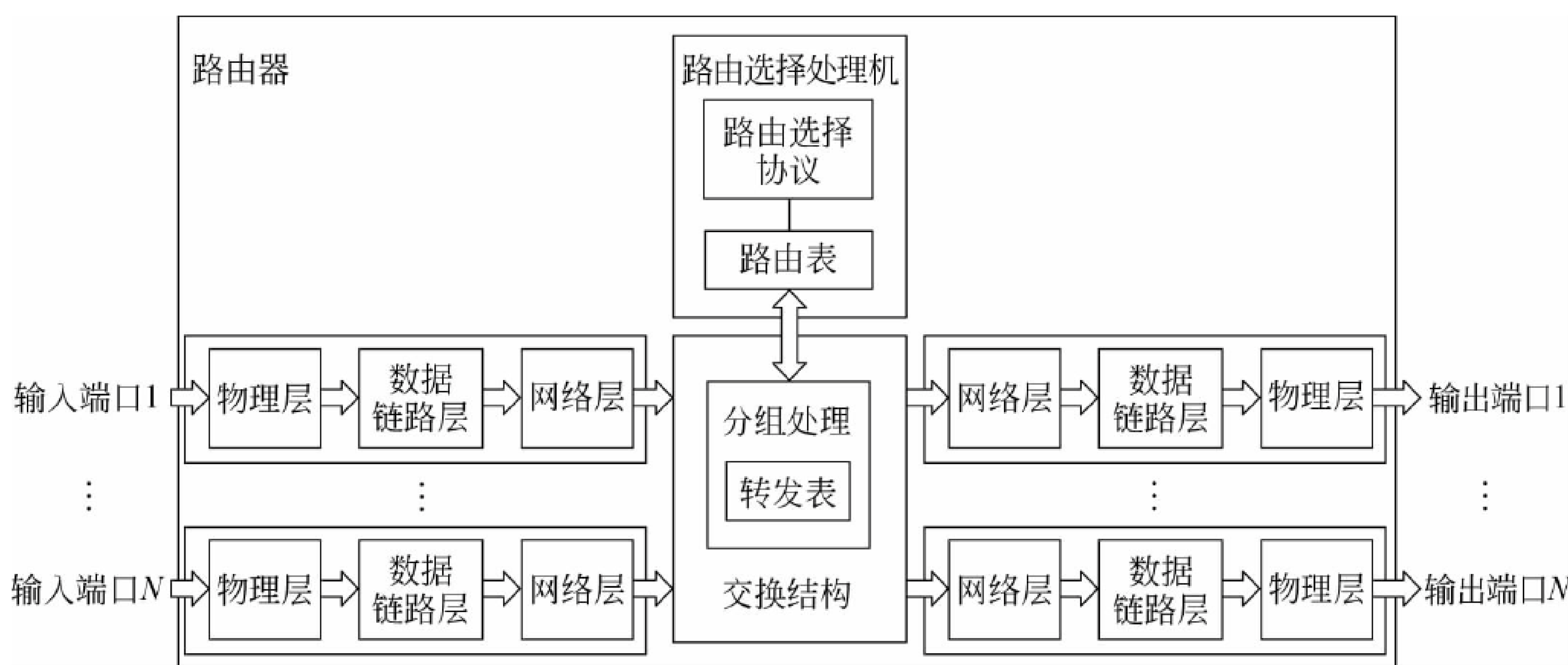


图 5-43　典型的路由器结构示意图

2）分组处理与交换部分

分组处理与交换部分主要包括交换结构、输入/输出端口。

(1) 交换结构

交换结构的作用是根据路由表和接收分组的目的 IP 地址,选择合适的输出端口转发出去。路由器是根据转发表转发分组,而转发表是根据路由表形成的。

(2) 输入/输出端口

路由器通常有多个输入端口和多个输出端口。每个输入和输出端口中各有三个模块,分别对应于物理层、数据链路层和网络层。物理层模块完成比特流的接收与发送;数据链路层模块完成拆帧和封装帧;网络层模块处理 IP 分组头。

如果接收的分组是路由器之间交换路由信息的分组(例如 RIP 或 OSPF 分组),则将这类分组送交路由器的路由选择处理机。如果接收到的是数据分组,则按照分组目的地址在转发表中查找,决定合适的输出端口。

典型的路由器外形结构如图 5-44 所示。

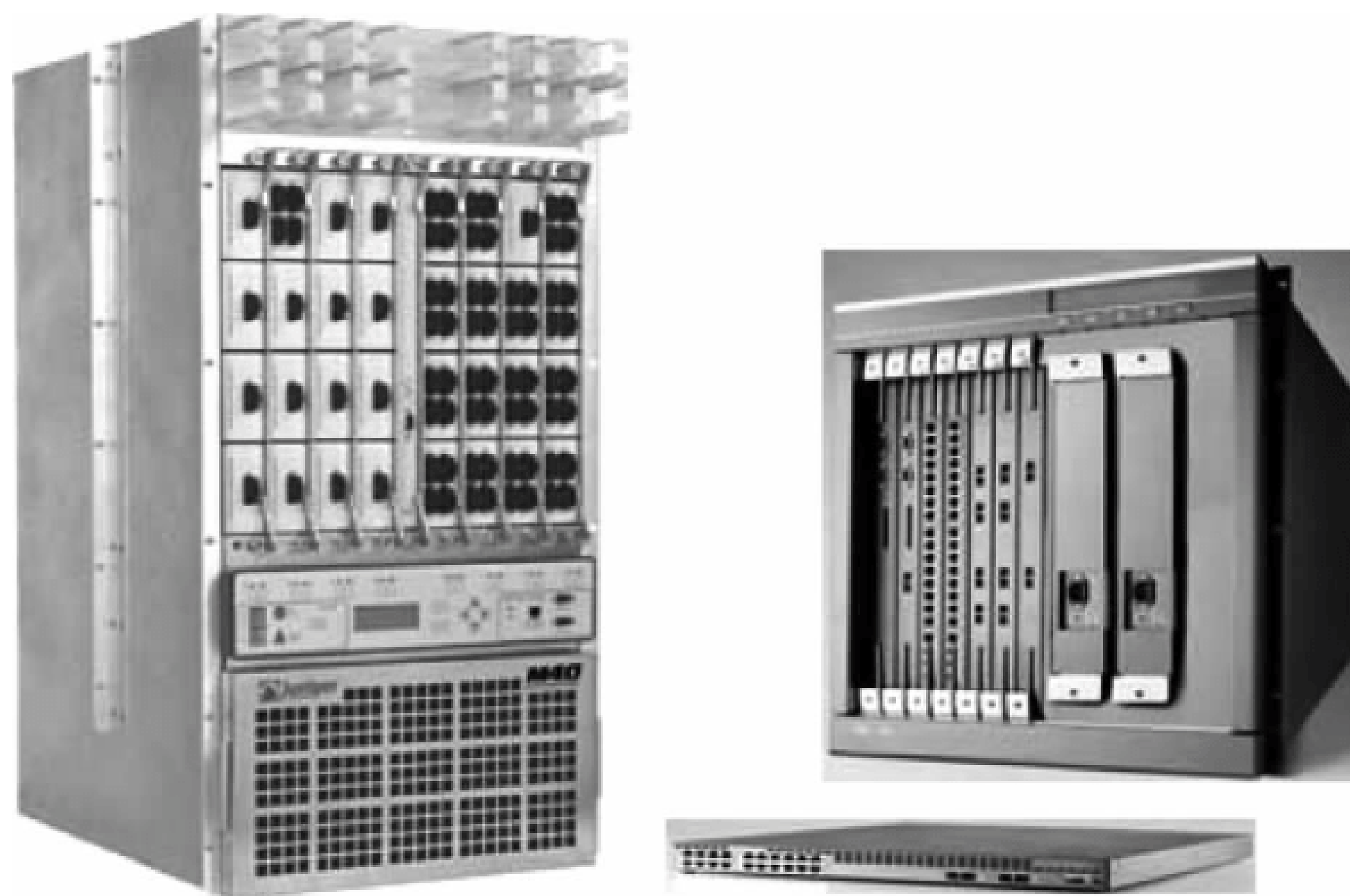


图 5-44　典型的路由器外形结构

2. 路由器的配置

路由器是互联网的核心设备。了解路由器的配置的基本概念,对于深入学习路由器工

作原理、网络设计与设备选型是很重要的。图5-45给出了典型的路由器配置所涉及的基本内容。

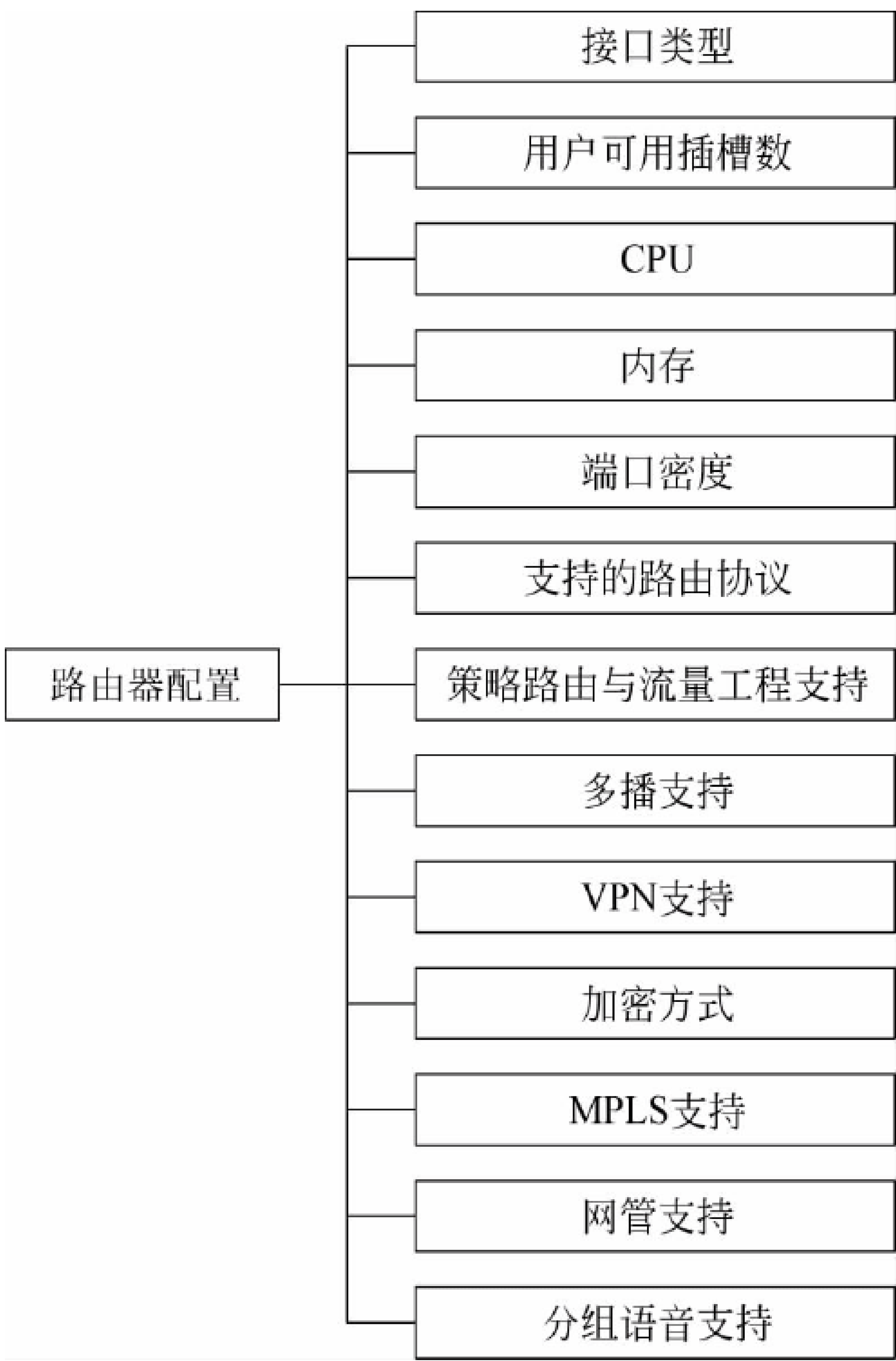


图5-45　路由器配置的基本内容

1）路由器支持的接口类型

路由器常见的接口类型包括Ethernet接口、串行接口、E1/T1、E3/T3接口，以及ATM与ISDN接口。Ethernet接口进一步分为10Mbps、100Mbps、1Gbps、10Gbps、100Gbps接口；串行接口进一步分为RS-232与RS-449DTE/DCE、X.21 DTE/DCE，高速E1同步串口支持的速率为2.048Mbps；POS接口进一步分为155Mbps、622Mbps、2.5Gbps、10Gbps与40Gbps接口。

2）用户可用插槽数

用户可用插槽数是指除路由器的CPU板卡、时钟板卡等系统板卡专用插槽之外，用户可以自行插入连接子网的线卡（Linecard）数量。可以根据用户可用插槽数，以及每一块线卡连接的端口数，来计算路由器所能够支持的最大端口数。

3）CPU

无论是高端路由器或者是中低端路由器，都要使用一个或多个CPU。CPU的处理能力决定了路由表的查找速度与路由计算速度，直接影响到路由器的吞吐率和转发延迟。

4）内存

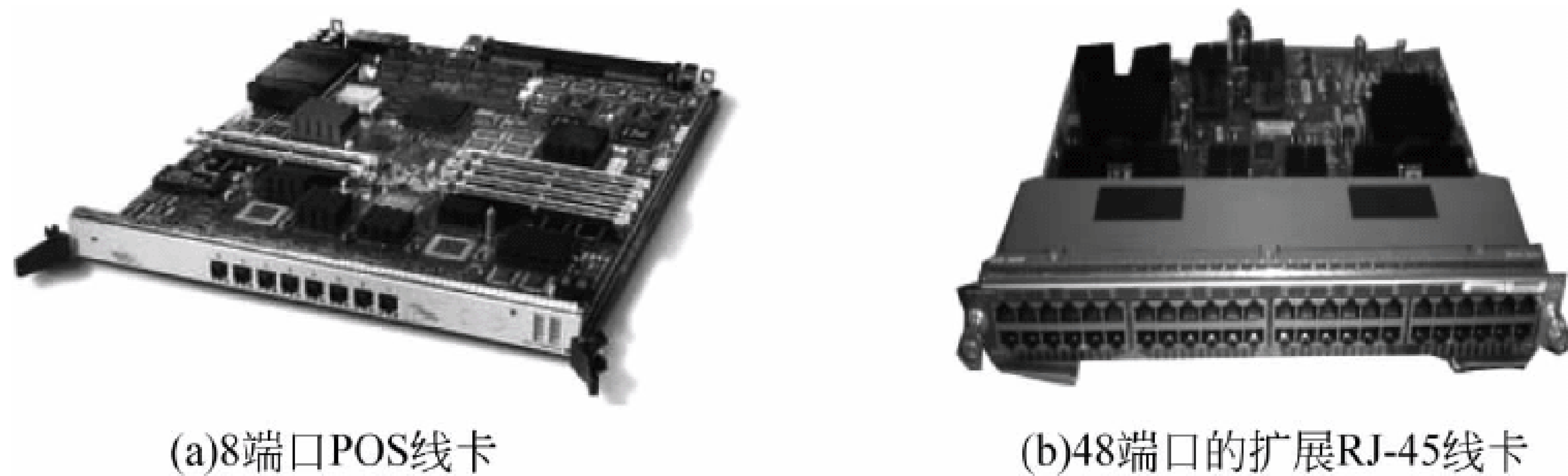
当路由器正在为一个接收分组查找转发表，准备转发时，后面跟着从这个输入端口可能连续收到多个分组，由于不能及时处理，后到的这些分组就必须在输入队列中排队等待处理。同样，输出端口从交换结构接收分组，然后将它们发送到路由器输出端口的线路上，也

要设有一个缓存来存储等待转发的分组。只要路由器的接收分组速率、处理分组速率、输出分组速率小于线速,无论是输入端口、处理分组过程与输出端口都会出现排队等待,产生分组转发延时,严重时会由于队列容量不够而溢出,造成分组丢失。

路由器会使用到不同类型的内存,如 Flash、DRAM 与 SRAM。存储器用来存放路由器的操作系统、配置文件、协议软件。中低端路由器的路由表、转发表可以存储在内存中,高端路由器的路由表存放在内存中,转发表则存放在线卡的存储器中。路由器内存的大小直接影响到路由表的查找速度与路由计算速度。

5) 端口密度

随着路由器连接节点的增多,路由器需要插入多块线卡,每一块线卡有多个外接端口。不同类型的路由器线卡的端口数不同。图 5-46(a)是一个 8 端口的 POS 线卡,图 5-46(b)是一个 48 端口的扩展 RJ-45 线卡。



(a)8端口POS线卡　　(b)48端口的扩展RJ-45线卡

图 5-46　线卡结构示意图

由于路由器体积与能提供的接入端口数相关。端口密度的指标是用来描述路由器制造的集成化程度的。端口密度应折合成机架内每英寸的端口数。通常情况下可以用路由器对每一种端口支持的最大数量来代替。

6) 支持的网关协议

路由器支持的路由协议类型主要包括 RIP、OSPF、IS-IS、BGP、IEEE 802.3/802.1Q、IPv4 或 IPv6、源地址路由、透明网桥、PPP、MLPPP、PPPOE 等。其中,源地址路由是策略路由中的一种,一般路由器都应该支持;透明网桥是指路由器端口以透明网桥的方式工作,执行网桥的功能,只做 MAC 帧的桥接,不对分组做路由检查与转发;早期的网络是由路由器使用点-点方式连接起来的,并且大多数用户都使用 PPP 接入路由器,凡是采用串口的路由器都应该支持 PPP,并作为首选;MLPPP 是指将多个 PPP 链路捆绑使用。

7) 策略路由与流量工程支持

路由器除了将目的 IP 地址作为路由选择的依据之外,还可以根据 TOS 字段、源地址路由、源端口号与目的端口号来为分组选择路由。策略路由可以在一定程度上实现流量工程,使不同服务质量的数据流,以及实时语音与 FTP 等不同性质的数据按不同的路径传输。路由器可以通过选择流量工程(TE)或虚拟专网(VPN)方式来实现对 MPLS 功能的支持。

8) 支持多播协议

随着网络会议、视频播放、计算机协同工作(CSCW)的发展,越来越多的应用需要路由器支持多播服务。在这种应用中,路由器需要支持互联网多播组管理协议(IGMP)、距离矢量多播路由协议 DVMRP,以及协议无关多播(PIM)协议。

9）支持 VPN

很多网络应用都需要路由器支持 VPN 服务，这就需要路由器支持 MPLS VPN 协议、第二层隧道协议（Layer 2 Tunneling Protocol，L2TP）、通用路由封装协议（Generic Routing Encapsulation，GRE），以及 IP over IP 协议等。

10）加密方式

为了解决 IP 协议的安全性问题，IETF 提出了 IPSec 协议。IPSec 不是一个简单的协议，而是一个协议包。它是由三个主要的协议与加密、认证算法构成了 IPSec 安全体系。对于支持 IPv4 协议的路由器，IPSec 是可选的；对于支持 IPv6 协议的路由器，IPSec 协议是必选的。

11）支持网管

路由器必须支持网络管理员通过 SNMP 网络管理程序与基于 Web 的网管，对路由器进行配置管理、性能管理、安全管理、差错管理与记账管理。网络管理程序可以管理到路由器的端口、网段、IP 地址与 MAC 地址。目前有一些路由器集成了 NAT、防火墙等功能。

12）支持分组语音的能力

在企业与办公环境的应用中，支持分组语音的 IP 电话能力十分重要。IP 分组语音功能需要支持 H.323 标准、G.723 与 G.729 标准。路由器支持 IP 电话的能力一般是以 E1 计算。一个 E1 信道支持 30 路电话。中低端路由器一般能支持 DSS1 与中国 1 号信令。

问题 5-53：评价路由器性能的指标主要有哪些？

评价路由器的性能指标可以分为转发性能、QoS 能力、安全能力、可靠性与可用性指标 4 个方面。

1. 路由器的转发性能

路由器转发性能主要包括线速转发能力、吞吐量、背靠背帧数、丢包率、延时、延时抖动等。

1）全双工线速转发能力

全双工线速转发能力是指：路由器端口以最大速率双向接收与发送分组时，没有出现丢包。线速转发是评价路由器性能的一个重要指标。所谓端口的线速转发就意味着：从路由器的接收端口接收了多少个分组，就能够通过发送端口转发出多少个分组，不会由于路由器处理能力的限制而造成分组丢失。全双工线速转发能力的计算是在转发最小报文长度（Ethernet 端口 64B、POS 端口 40B）和最小包间隔（符合协议规定）的条件下进行的。因为以最小包间隔，接收和转发最小长度报文是对路由器转发性能最苛刻的考验。全双工线速转发能力的单位是 pps（packet per second）。

2）吞吐量

吞吐量分为端口吞吐量与整机吞吐量。端口吞吐量是指：路由器某一个端口每一秒钟转发的最大分组数。对于不同的线卡、同一个线卡的不同端口，端口吞吐量的实际测量值可能是不同的。

设备吞吐量是指：路由器整机在不丢包的情况下，每秒钟最多能够转发的最大分组数。设备吞吐量与路由器的端口数量、端口速率、分组长度、分组类型、路由计算模式（集中或分布），以及测量方法相关。路由器的设备吞吐量通常小于所有端口吞吐量的总和。

在测试路由器吞吐量时，通常是以一定速率向路由器发送一定数量与相同长度的报文，

并检测路由器在不丢失报文的前提下,能够转发最多的报文数。如果路由器接收的报文与转发的报文数量相等,那么就将发送速率提高,并重新测试。如果路由器转发的报文少于接收的报文数,则降低发送速率,重新测试,直至得出最终结果。吞吐量的单位可以使用pps或bps。全双工线速转发能力表示的是一种理想状态下路由器的转发能力,而吞吐量是路由器实际能够达到的报文转发能力。

3) 背靠背帧数

背靠背(back-to-back)帧数是指:在不丢失报文的前提下,以最小帧间隔发送的最多分组数。背靠背帧数指标是用来测试路由器的缓存能力。

4) 路由表能力

路由表能力是指:路由表能够存储的路由表项数量的多少。一般高端路由器应该能够支持250 000条路由。

5) 背板能力

背板是路由器输入端与输出端之间的物理通道。低端路由器采用共享背板的结构,中高端路由器一般采用的是交换结构。背板能力决定了路由器的吞吐量。

6) 丢包率

丢包率是指:在稳定状态和持续负荷的状态下,由于资源缺少而造成应该转发的分组,不能转发分组所占的比例。丢包率是衡量路由器在超负荷工作状态下的性能指标。

7) 延时

延时是指:分组的第一个比特进入路由器,到最后一个比特从路由器输出的时间间隔。延时表示存储转发工作模式中,路由器对分组处理所需要的时间。延时与报文长度、链路传输速率相关。路由器的延时通常是在路由器端口吞吐量的测量范围内测试,它对网络性能影响很大。

8) 延时抖动

延时抖动是指:在正常工作状态下,延时的变化量。一般情况下,数据传输对延时抖动不敏感,而在语音与视频传输中对延时抖动要求很高。

2. QoS能力

路由器保证服务质量的QoS主要包括:队列管理机制、端口硬件队列数、QoS分类方式、分类业务带宽保证、资源预留RSVP与区分服务DiffServ支持、承诺接入速率等。

队列管理机制通常是指:路由器拥塞管理机制与队列调度算法,如随机早期检测(Random Early Detection,RED)、加权随机早期检测(Weighted Random Early Detection,WRED)、加权循环(Weighted Round Robin,WRR)、加权公平队列算法(Weighted Fair Queuing,WFQ)等队列调度算法。

路由器支持的优先级是由端口硬件队列来保证的。每个队列中的优先级是由端口硬件队列调度算法来保证的。因此,端口硬件队列数与队列调度算法对路由器QoS保证指标的实现是重要的。

路由器对服务质量QoS的支持所依据的信息是不同的。最基本的QoS分类是基于端口,也可以是基于802.1Q中规定的数据链路层,以及网络层IP报头中的TOS字段、源与目的地址,高层的源与目的端口号。

分类业务带宽保证体现在路由器是否对不同业务等级做出带宽保证,由队列调度算法

等方法实现。

目前，在IP协议基础上支持QoS服务的协议主要有资源预留（RSVP）、区分服务（DiffServ）与多协议标记交换（MPLS）。目前，很多路由器都支持MPLS协议。

承诺接入速率（Committed Access Rate，CAR）或者叫作速率限制，是对特定种类通信的带宽进行控制，以保证这些流量不会影响重要的通信流量。一方面，CAR可以起到流量控制的作用；另一方面，可以通过在接入路由器上设定CAR值，通过限速策略达到抵御DoS攻击的目的。因为DoS/DDoS攻击的一个重要特征是网络中会出现大量带有非法源地址的ICMP报文，通过在路由器上对ICMP报文配置CAR，来设置速率上限的方法来及时中断DoS/DDoS攻击。

3. 安全性

了解路由器的安全性指标，可以从支持VPN、访问控制能力与对网络流量过滤的能力入手。

通常路由器都应该具有支持VPN的能力。路由器产品支持VPN的功能差异表现在支持哪种标准的VPN，例如支持IPSec VPN，以及最多可建立几条VPN安全隧道。

一般的路由器访问控制能力表现在通过数据分组的源IP地址、源端口、目的IP地址、目的端口以及设定的时间段，控制网络内部主机对外网的访问。访问控制能力较强的路由器还可以通过定义工作组、服务端口、协议、时间段，实现对网络内部用户对外部网络的访问，或网络外部用户对内部网络设备与用户的访问；通过IP地址/MAC地址的绑定功能，实现用户身份的识别，防止IP地址盗用、MAC地址盗用以及IP/MAC欺骗等常见攻击。

路由器可以采用对特定报文流量限制，如对ICMP报文配置CAR，通过设置速率上限的方法，来及时地发现和中止DoS/DDoS等攻击。

4. 可靠性与可用性

路由器的冗余的目的是保证设备的可靠性与可用性。对于路由器来说，冗余应该包括接口线卡冗余、电源冗余、系统板冗余、时钟板冗余与设备冗余。至于需要采用哪一类冗余，需要在设备可靠性与投资之间取得平衡。

高性能路由器的设计中注意"多级冗余保护，避免单点故障"，实行板卡的冗余，有故障的交换链路、控制卡、交换卡都不影响分组转发。同时考虑将一个接口线卡与多个交换卡互连，实现数据交换网络内部连接的冗余。硬件检测机制快速检查出故障之后，可以通过冗余连接的线路绕过故障交换卡。Tbps的路由器采用了超立方、三维环等多级交换网络，使用了更多的冗余路径。

路由器冗余需要RFC2338的虚拟路由器冗余协议（Virtual Router Redundancy Protocol，VRRP）来支持。

路由器需要24小时连续工作，因此更换部件应该不影响或尽量减少对路由器可用性的影响。热插拔是指在不影响系统运行的情况下，对出现问题的模块进行插拔操作。路由器选型是需要了解该型号的路由器是否支持部件的热插拔，以及支持哪些部件的热插拔。

使用POS端口的路由器互连时需要考虑路由器与POS（Packet-Over-SONET）设备的同步问题，因此需要按照POS设备的要求，注意路由器内部时钟的精度与配置方法。

路由器的可用性可以用无故障连续工作时间、系统故障恢复时间、主备用系统切换时间与SDH接口自动保护切换时间来描述。路由器的可用性是网络管理人员最关心的指标之

一。计算路由器的可用性指标的公式是:

可用性=MTBF/(MTBF+MTBR)

其中,MTBF为平均无故障时间,是指两次故障之间时间间隔的平均值,它反映了路由器连续正常运行的时间长度;MTBR为平均故障修复时间,是指两次故障修复时间的平均值,它反映了路由器故障恢复能力。

如果路由器的可用性达到99.9%,那么每年停机时间≤8.8h,可用性达到99.99%,那么每年停机时间≤53min,可用性达到99.999%,那么每年停机时间≤5min。

理解路由器性能指标时,需要注意以下几个问题。在实际应用中,高端路由器的可靠性与可用性指标选型时一般要控制在:

(1) 系统可用性≥99.999%;

(2) 系统无故障连续工作时间MTBF>10万小时;

(3) 系统故障恢复时间<30min;

(4) 系统应具有自动保护功能,主备用切换时间<50ms;

(5) SDH与ATM接口应具有自动保护功能,主备用切换时间<50ms。

要求设备具有高可靠性与高稳定性。主处理器、主存储器、交换矩阵、电源、总线仲裁器与管理接口等系统主要部件应具有热备份冗余。线卡要求$m+n$备份,并提供远端测试诊断功能。

问题5-54:如何认识路由器的发展趋势?

最初的简单路由器可以是由一台普通的计算机加载特定的软件,并增加一定数量的网络接口卡构成。特定的软件主要实现路由选择、分组接收和转发功能。为了满足网络规模发展的需要,高性能、高吞吐量与低成本的路由器的研究、开发与应用,一直是网络设备制造商与学术界十分关注的问题。

在讨论路由器发展背景时,需要注意两个信息技术领域的预测定律。一是“摩尔定律”。根据摩尔定律,半导体器件的计算能力大约每18个月翻一番。那么,路由器作为一种特殊结构和用途的计算机,其计算能力也应该按照这个速度增长。二是“吉尔德定律”。根据吉尔德预测:主干网的带宽将每6个月增加一倍。显然,受到半导体器件的计算能力限制的路由器的计算能力的增长,要远远低于互联网网络带宽的增长速度。目前,光纤通信技术的快速发展,使得作为互联网主干网主要传输介质的光纤带宽与路由器计算能力相比,光纤带宽已经不是限制网络性能的瓶颈。传统的路由器已经成为互联网网络带宽增长的瓶颈。如果仍然采用传统的路由器体系结构,那么按照“摩尔定律”增长的半导体器件的计算能力无法满足实际应用的需求。因此,如果要从根本上解决路由器的瓶颈问题,出路在设计新的路由器体系结构上。

随着Internet的广泛应用,路由器的体系结构经历了几次大的变革。这种变化的特点主要集中在以下两个方向。

(1) 从基于通用功能芯片的结构向使用专用芯片结构的方向发展。

(2) 从基于系统的串行处理向并行处理的方向发展。

专门为路由器设计的专用芯片随着集成度的不断增长,性能也在不断提高。典型的路由器专用芯片是网络处理器(Network Processor,NP)。2013年9月12日Cisco公司发布的当前世界上性能最高的可编程网络处理器nPower,内部集成了40亿个晶体管。高性能、

可编程网络处理器 nPower，使得路由器的一些关键运算可以采用独立的硬件来完成。路由器系统可以通过针对路由运算特点，对算法进行优化设计，并采用多个处理单元并行处理，因此可以大幅度地提高路由器的系统性能。

路由器的体系结构大致经历了以下 4 次较大的变迁。

（1）单总线单 CPU 结构的路由器；

（2）多总线多 CPU 结构路由器；

（3）交换结构的 Gbps 路由器；

（4）第三层交换与多级交换路由器。

1. 单总线单 CPU 结构的路由器

最初的路由器采用了传统的计算机体系结构，它包括 CPU、内存 RAM 和挂在总线 BUS 上的多个连接网络的接口卡。采用单总线单 CPU 结构的路由器与一台通用的计算机没有很大的区别。1986 年，由 Cisco 生产出世界上第一台路由器产品 AGS（Advanced Gateway Server），其外形与结构如图 5-47 所示。



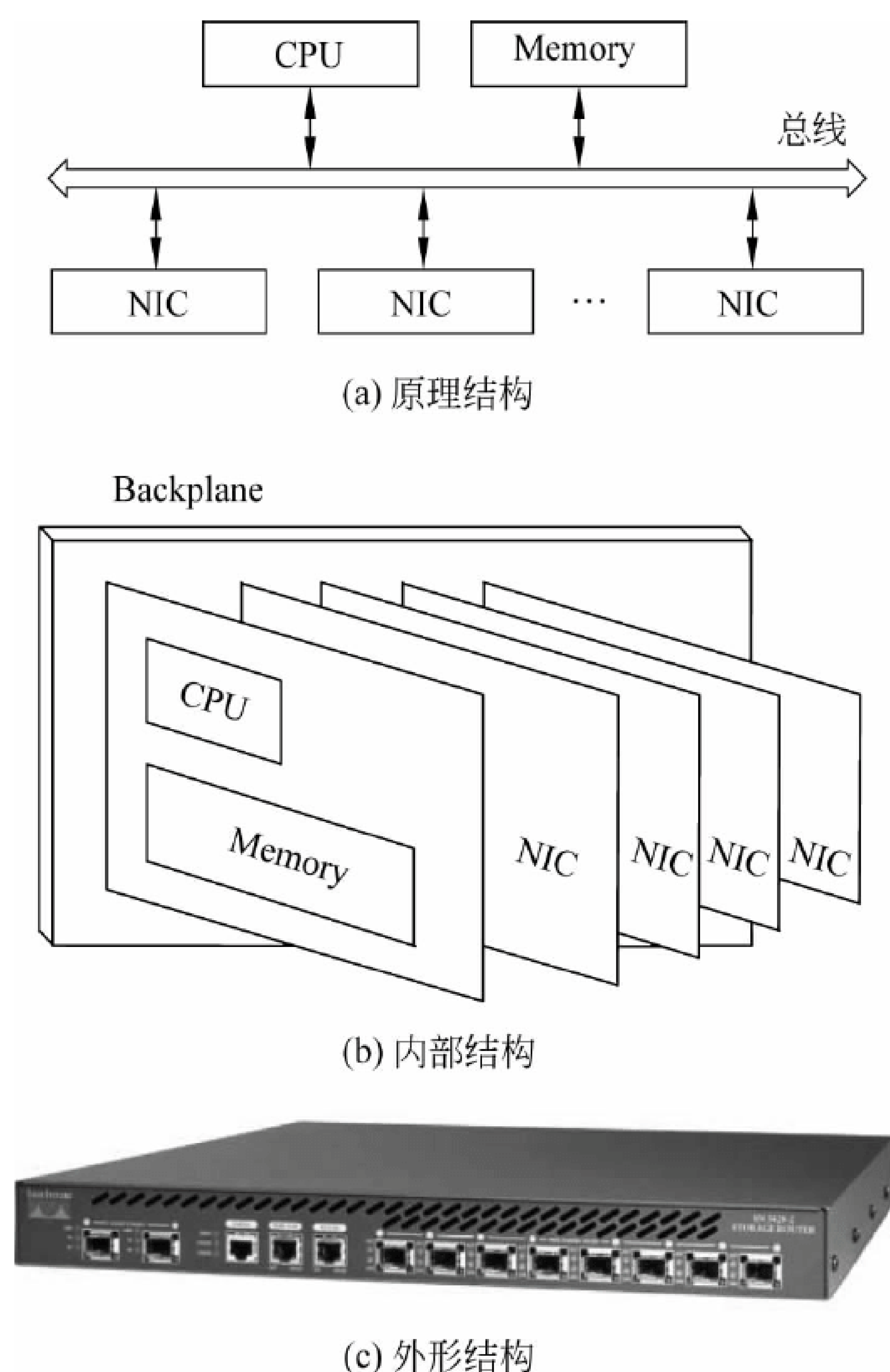
图 5-47　世界上第一台路由器产品

Cisco AGS 采用的就是单总线单 CPU 的结构，其中，CPU 用的是 Motorola 的 MC68020 处理器。图 5-48(a)给出了 Cisco AGS 的原理结构，图 5-48(b)给出了 AGS 的内部结构，图 5-48(c)给出了 AGS 的外形结构。

在讨论路由器结构时，需要注意以下几个问题。

（1）从原理示意图中可以看出，在 AGS 路由器中 CPU 担负着检查 IP 分组头、计算校验和与生存时间 TTL，以及计算路由和维护路由表的功能。网卡担负着 IP 分组的接收与发送的功能。

（2）从内部结构示意图中可以看出，CPU 与 Memory 的主板与所有的网卡 NIC 都是插在路由器背板的扩展槽中。背板中的总线负责路由器几种板卡之间的信息交互。路由器采用这种结构的好处在于：当接入的子网采用协议发生变化，网卡 NIC 改变后，或者需要增减网卡 NIC 时，可以通过简单的拔插网卡的方式来实现。背板的总线带宽对路由器的吞吐率等性能参数有很大的影响。



(a) 原理结构
(b) 内部结构
(c) 外形结构

图 5-48　单总线单 CPU 结构的路由器结构示意图

（3）由于 AGS 路由器采用的是单总线单 CPU 的结构，当网卡 NIC 从与它连接的网络中接收到 IP 分组时，NIC 通过单总线将数据存放在内存之中，同时通知 CPU 对接收到的 IP 分组头进行处理。CPU 在判断接收到的分组正确之后，查找路由表，决定 IP 分组由哪一个网卡 NIC 转发。负责转发该 IP 分组的网卡 NIC 通过 MAC 层、物理层，将它转发到下一个路由器。

（4）网卡 NIC 与 CPU、存储器，通过单总线的多次交互，才能完成一个分组的转发。当路由器背板插入的 NIC 增多时，不同板卡之间对总线、CPU 的争用会引起严重的访问冲突，导致系统处理能力的下降。因此，单总线单 CPU 结构的路由器存在着处理速度慢，CPU 的故障将导致系统瘫痪的缺点。

2. 多总线多 CPU 结构路由器

为了提高路由器的性能，出现了多总线多 CPU 结构的路由器。多总线多 CPU 结构的路由器可以进一步分为三种类型：单总线主从 CPU 结构、单总线对称多 CPU 结构、多总线多 CPU 结构。

1）单总线主从 CPU 结构

第一种是单总线主从 CPU 结构的路由器，两个 CPU 是非对称的主从式结构关系，一个 CPU 负责数据链路层的协议处理，另一个 CPU 负责网络层的协议处理。典型的产品有 3COM 公司的 Net Builder2 路由器。这种路由器是第一代的单总线单 CPU 结构的简单延伸。路由器的系统容错能力有比较大的提高，但是数据包的转发处理速度并没有明显提高。

2）单总线对称多 CPU 结构

针对单总线主从 CPU 结构的缺点，第二种单总线对称多 CPU 结构中开始采用并行处理技术。在每个网络接口处使用一个独立的 CPU，负责接收和转发本接口的数据包，其中包括队列管理、查询路由表和决定转发。主控 CPU 则完成路由器的配置、控制与管理等非实时任务。典型的产品有 Bay 公司的 BCN 系列路由器，它的 CPU 使用的是 Motorola 的 MC68060 和 MC68040 处理器。尽管这种结构的路由器的网络接口处理能力提高，但是单总线与软件实现转发处理这两个因素成为限制路由器性能提高的瓶颈。

3）多总线多 CPU 结构

针对单总线与软件转发处理的问题，第三种路由器将多总线多 CPU 结构与路由加交换技术相结合。多总线多 CPU 结构的路由器至少需要使用三种或三种以上的 CPU 与总线。

Cisco 7000 系列的路由器是一种典型的多总线多 CPU 结构的产品。Cisco 7000 系列的路由器使用三种 CPU 与三种总线。三种 CPU 是接口 CPU、交换 CPU 和路由 CPU，三种总线是 CxBUS、dBUS 与 SxBUS。

图 5-49(a)给出了多总线多 CPU 的路由器原理结构，图 5-49(b)给出了多总线多 CPU 的路由器外形结构。多总线多 CPU 结构的路由器也称为“单机分布式总线结构路由器”。

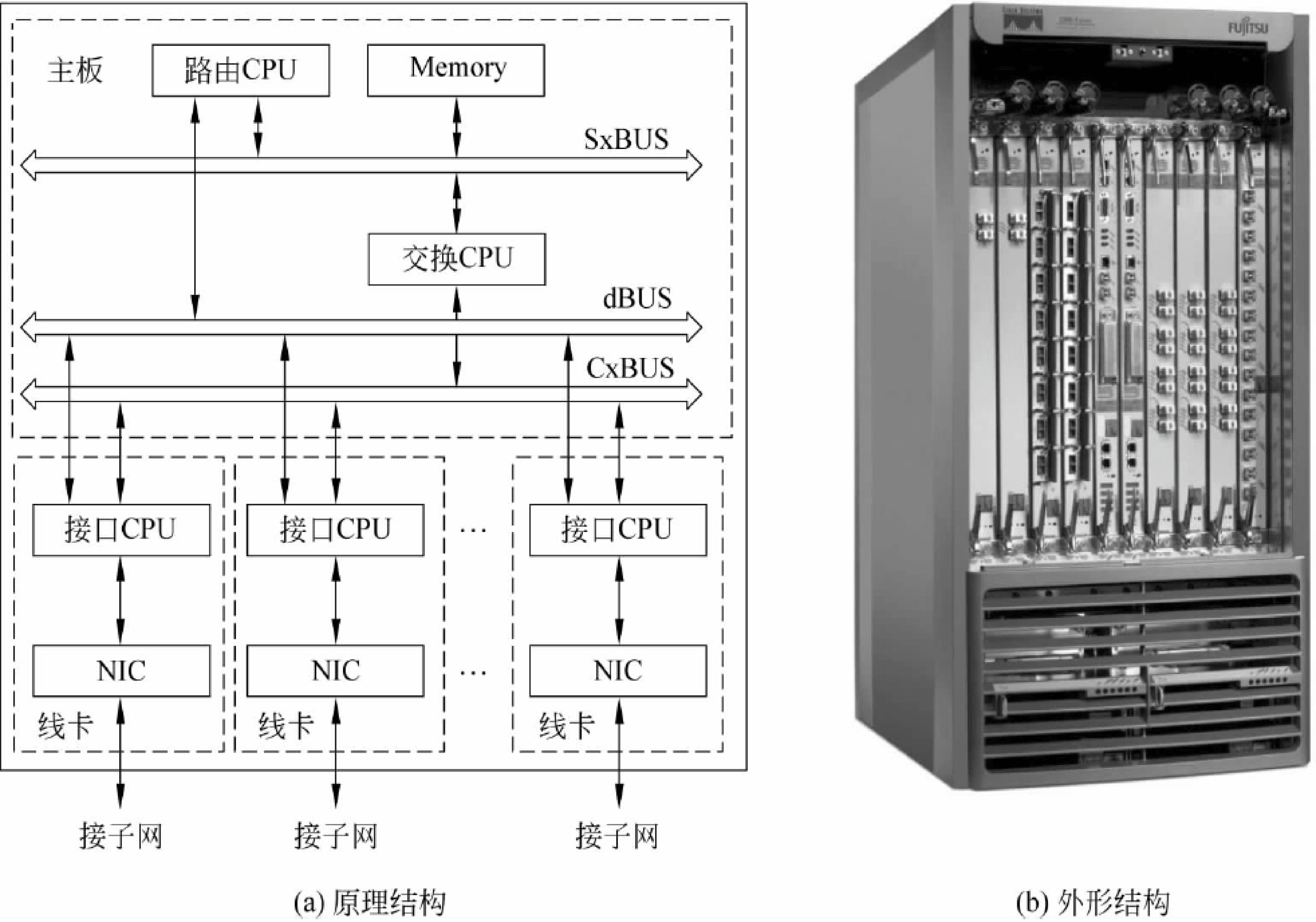


图 5-49　多总线多 CPU 路由器结构示意图

讨论多总线多 CPU 结构路由器特点时，需要注意以下几个问题。

(1) 多总线多 CPU 结构路由器的内部结构主要是由主板、线卡组成。一台路由器可以插入多块线卡。线卡(Linecard)是线路接口子系统，它为路由器提供多种类型的外部接口，实现路由器与外部子网的连接。

路由器的每一块线卡都有自己的接口 CPU、内存与网卡 NIC，组成能够独立处理 IP 分

组的子系统。多块线卡可以并行处理接收不同子网的IP分组,将接收到的分组存储在自己的内存中。接口CPU在判断分组接收正确,并根据目的IP地址进行路由表查找之后,再通过分组传输到对应的线卡。

(2) 主板基本上不参与路由转发操作。主板主要执行网关协议,与邻居路由器交换路由信息,生成、更新和维护各线卡存储的路由表,而分组转发操作分布在各线卡中进行。在路由与交换技术方面,系统采用硬件Cache快速进行路由表查找,以提高转发处理的速度。

(3) 在使用单总线的路由器中,多个网卡NIC争用共享总线时出现冲突是造成路由器性能下降的重要原因。尽管多总线多CPU结构路由器中增加了不同用途的总线,但是多个线卡共享一条总线的局面仍然存在,在多总线多CPU结构中共享总线仍然是提高路由器性能的瓶颈,因此理论上多总线多CPU结构路由器的最高转发能力可以达到5Gbps。1993年Cisco公司生产的7000系列的路由器最高转发能力达到2Gbps。它是第一款用作IP主干网的多总线多CPU结构路由器。

3. 基于交换结构的路由器

1) 基于硬件交换路由器的基本设计思想

研究人员发现,借用传统的计算机结构设计思想、软件交换的方式,无法实现路由器端口10Gbps或2.5Gbps的线速转发,必须在设计思想上加以改进,采用基于硬件的交换结构,去替代传统路由器共享总线的软件交换结构。基于硬件交换的路由器结构可以有三种设计思路:基于内存交换、基于总线交换与基于交叉结构。图5-50给出了基于硬件交换的三种路由器结构。

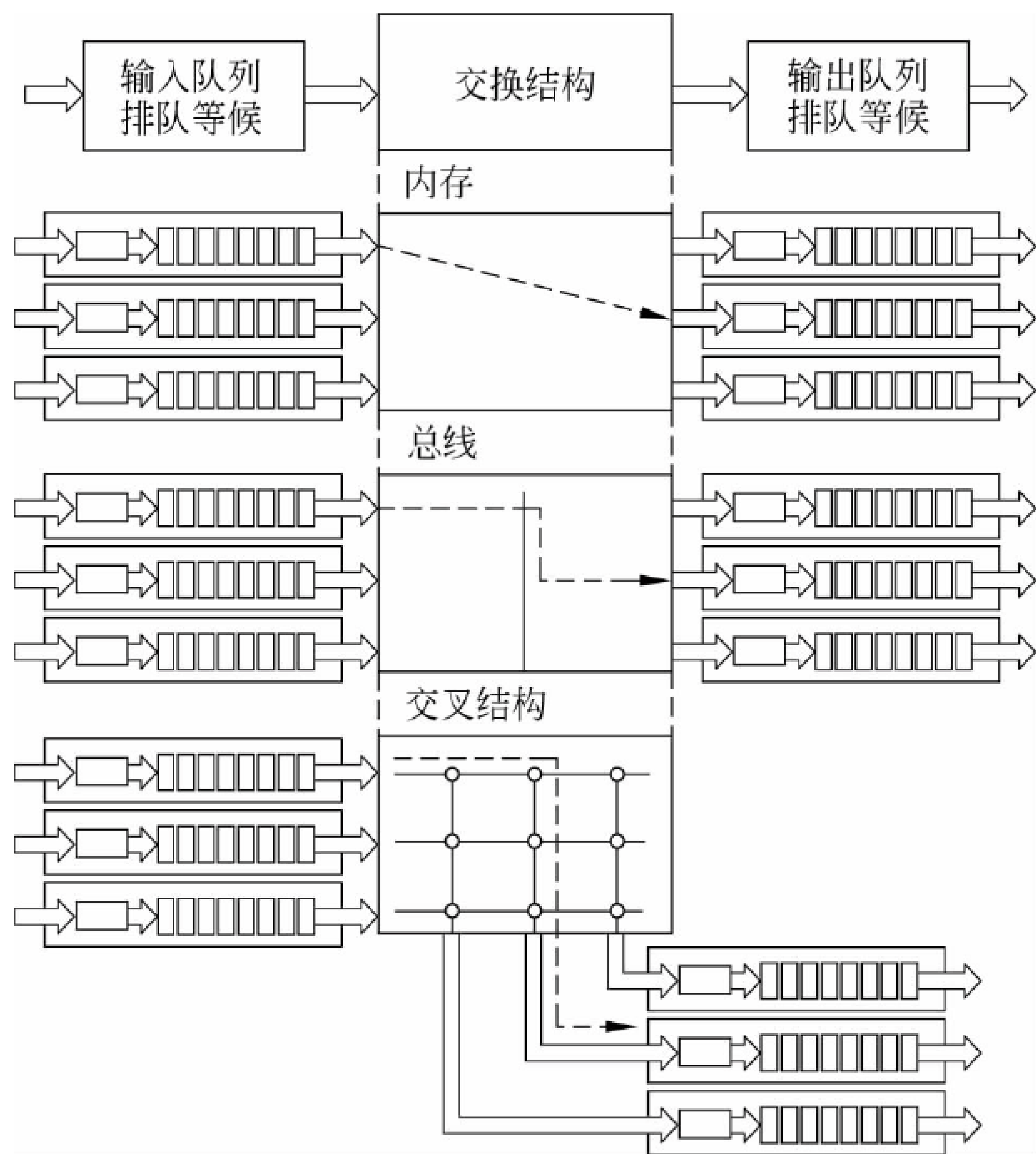


图5-50 基于硬件交换的路由器结构

2) 基于交叉结构的路由器

在三种方案中,基于交叉结构的设计方案是最优化的。交叉结构又称为“交换结构”,它

是通过专用大规模集成电路 ASIC 实现的“交叉开关(Crossbar)”与相关电路,完成多路数据的并发交换。图 5-51 给出了基于交叉结构的路由器结构。

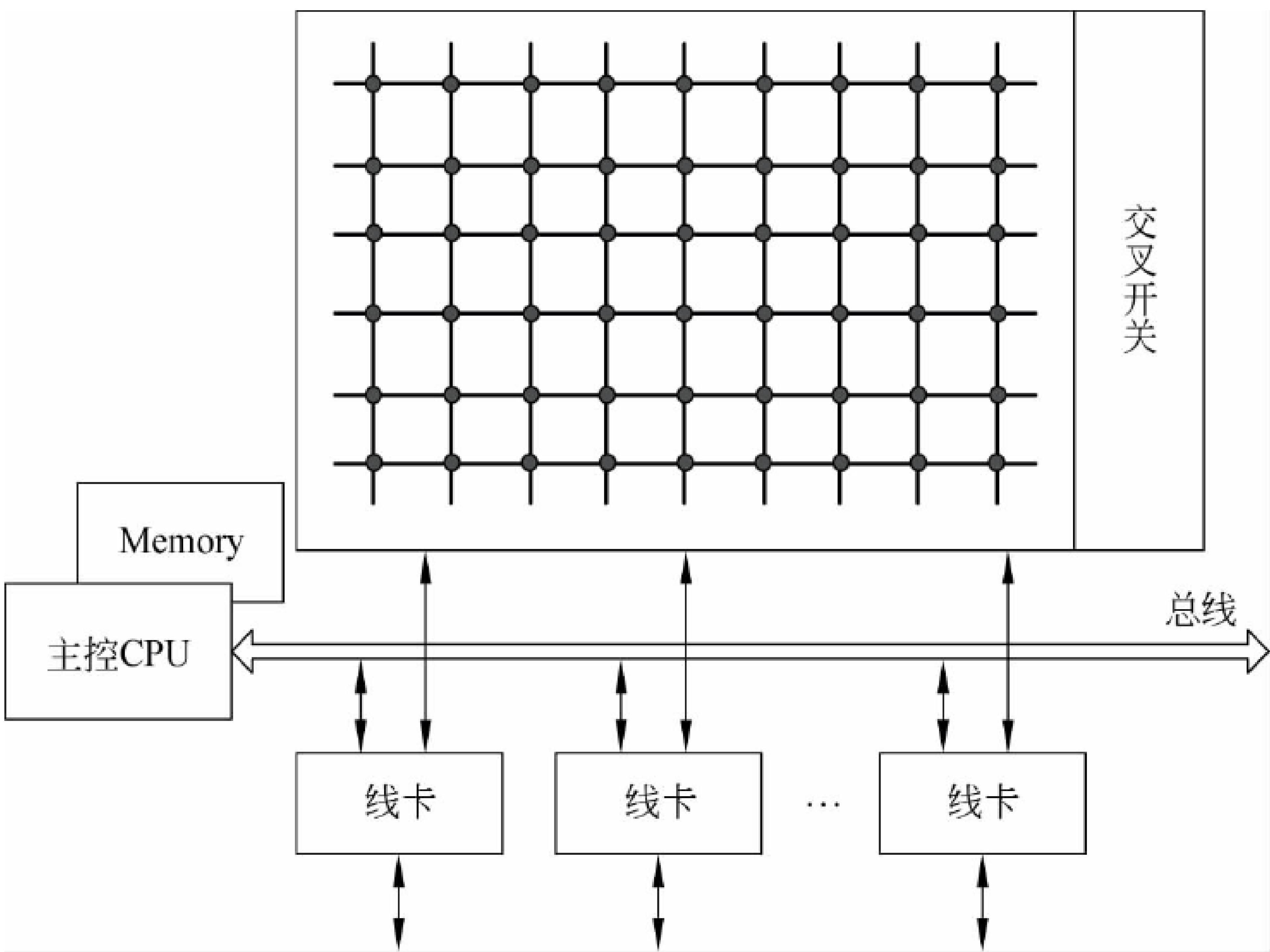


图 5-51　基于交叉结构的路由器结构

在基于交叉结构的路由器中,交叉开关代替了共享总线。由于各个线卡接收到 IP 分组后可以并行、独立地处理,然后分别通过交叉开关中不同的线路,直接传送到输出的线卡,因此可以有效地提高路由器的吞吐率,降低转发延时。这样,路由器的性能直接由中心交叉开关与各板卡的性能决定,而不是仅取决于总线的带宽。同时,由于路由器省去了大量的存储器模块,因此可以减少路由器系统结构的复杂性,降低路由器的成本。

基于交叉结构路由器的典型产品有 Cisco 12000 路由器,最多能给支持 16 个 2.5Gbps 的 POS(Packet Over SONET/SDH)端口,可以实现多路数据的并发线速转发。由于该路由器没有集中的核心 CPU,所有线卡都有功能相同的 CPU,因此这种结构的路由器扩展性很好。路由与转发软件采用并行处理方法设计,可以有效地提高路由器的性能。基于交叉结构的路由器是核心路由器设备的首选类型。图 5-52 给出了基于交叉结构的 Gbps 路由器结构。

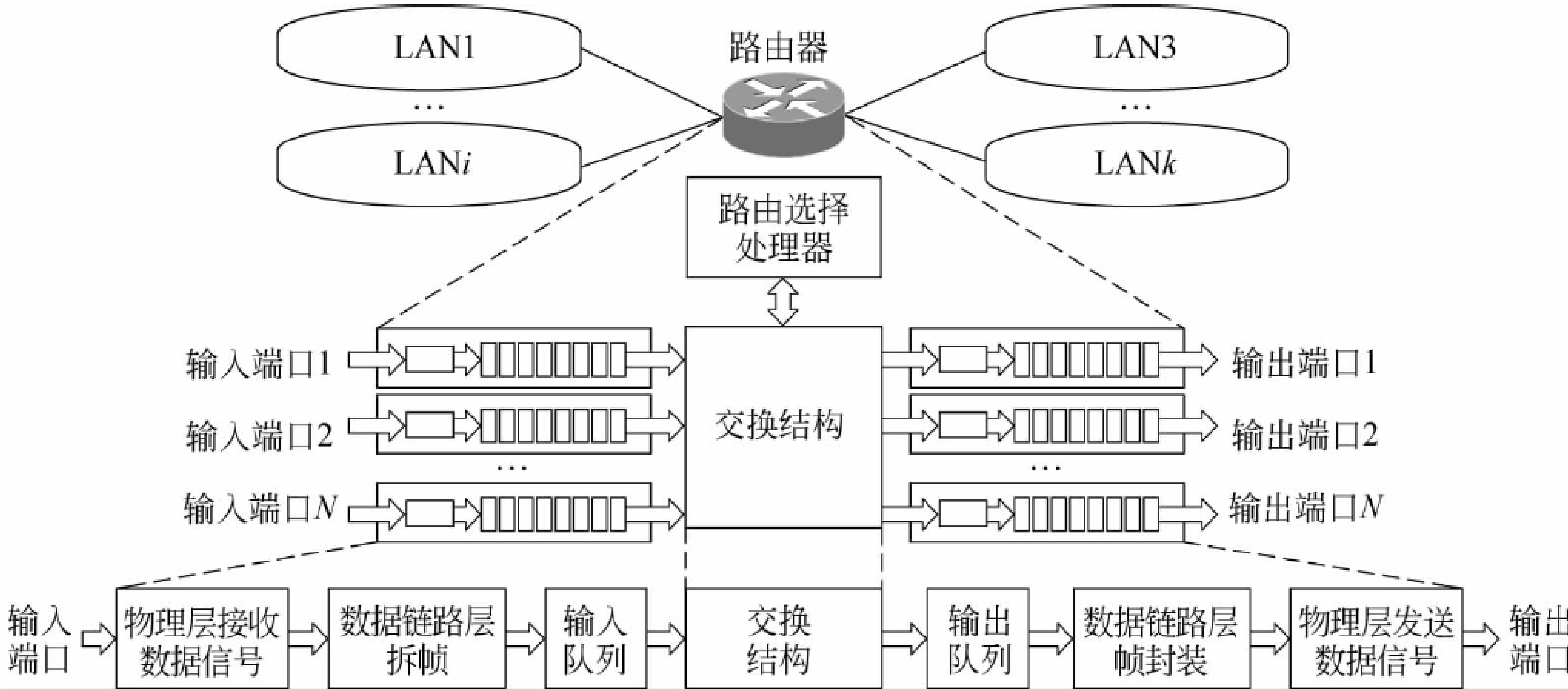


图 5-52　第三代路由器结构示意图

3）共享并行处理器的路由器结构

随着互联网大规模的应用，人们对路由器性能的要求也越来越高。在研发新的路由器时，人们会发现两个问题。第一，专用的ASIC芯片是路由器的成本增加，同时路由器生产厂家需要一只熟悉专用ASIC芯片的研发队伍。第二，由于互联网新的应用不断涌现，不同的网络应用促使网络协议与对网络性能要求的不断改变，而专用ASIC芯片的研发周期较长，不能适应网络应用快速发展的局面。在这样的背景之下，在路由器的设计思想上发生了两点重要的变化。

变化之一：网络处理器(NP)概念的提出。

NP是针对网络应用共性的需求，专门设计的一类适合开发网络应用的大规模集成电路VLSI芯片。NP采用了多微处理器(Multi-Microprocessors)的并行处理模式，具有很好的可编程能力。同时，网络设备制造商与网络处理器制造商共同提出了通用交换接口(Common Switch Interface，CSI)标准，使得NP成为一种标准的网络处理器件。典型的NP芯片产品如Intel公司的IXP系列、IBM公司的NP4GS3系列、MMC公司的np7000系列，以及EZchip的NP-2、NP-4系列的芯片。

路由器开发人员在掌握了NP的开发思路与并行软件编程方法的基础上，就可以通过灵活的软件编程，以较快的速度开发出适合不同性能要求的路由器产品，从而适应网络应用快速发展的需要。NP芯片可以用于路由器、防火墙、QoS控制与流量均衡设备之中。基于NP的共享处理器路由器结构如图5-53所示。

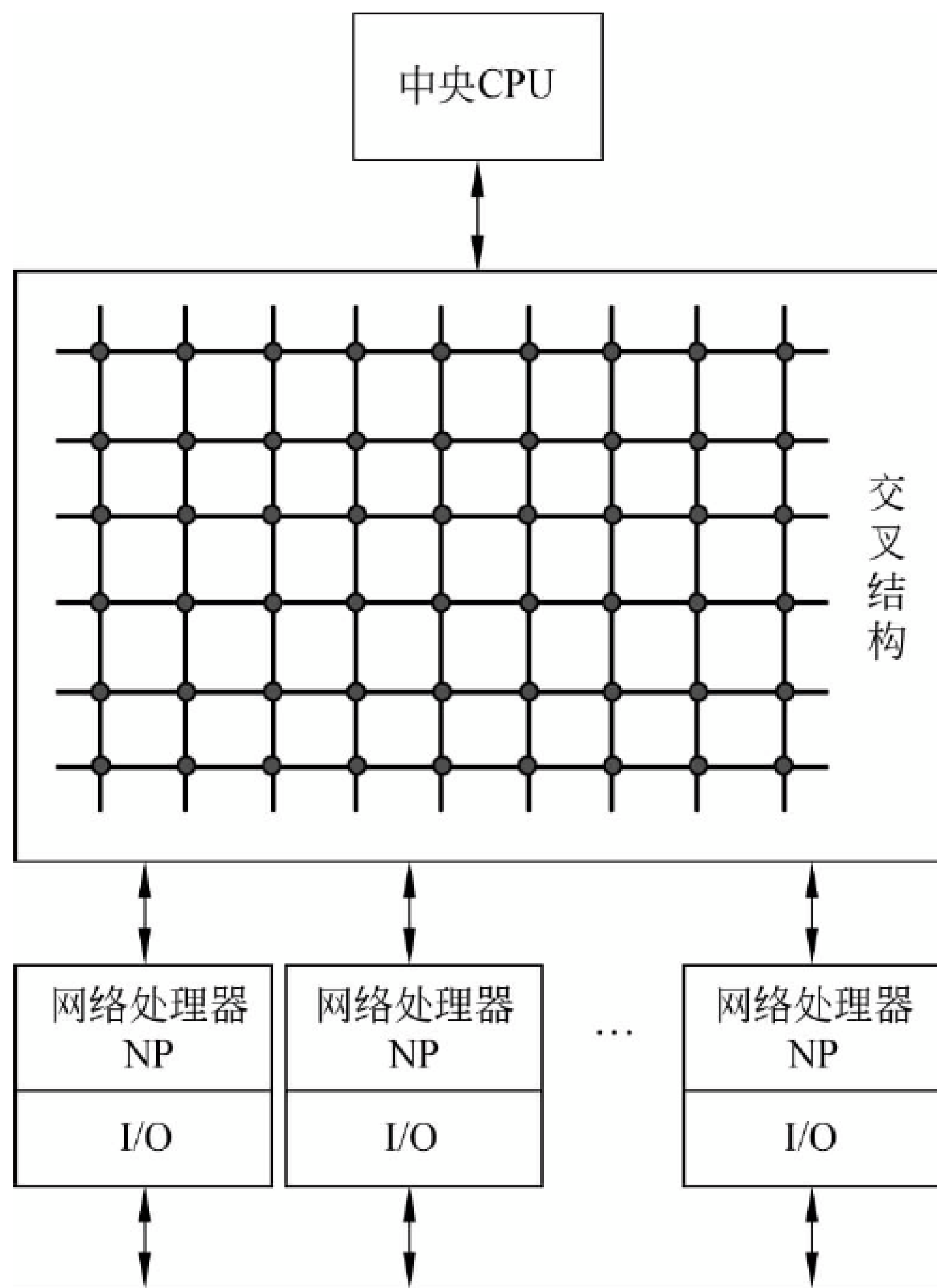


图5-53　基于NP的共享处理器路由器结构示意图

变化之二："第三层交换"概念的提出。

最初，人们将"第三层交换"的概念限制在网络层。但是，有一种发展趋势是：将第三层成熟的路由技术与第二层高性能的硬件交换技术相结合，可以达到快速转发，保证服务质量

(QoS),提高路由器性能的目的。

Ipsilon 公司最早开展将第三层路由与第二层交换结合的研究,并开发了 IP Switching 产品。随之其他公司也纷纷推出各自的产品,例如,Cisco 的标记交换 Tag Switching 产品、IBM 公司的汇聚基于路由的 IP 交换产品、Toshiba 公司的信元交换路由 CSR 产品等。这些产品都希望提高 IP 分组的转发速度,改善 IP 网络的吞吐量与延时特性。第三层交换机通过内部网关协议(例如 RIP 或 OSPF)创建和维护路由表。出于安全方面的考虑,第三层交换机通常提供防火墙分组过滤等服务功能。

4. 高可扩展路由器与路由器集群结构

随着互联网、移动互联网与物联网应用的大规模扩展,接入路由器的节点数、吞吐量 Tbps 级的快速增长,使得路由器的应用形态与体系结构发生了很大的变化。这种变化表现在两个方面,一个是高可扩展路由器结构;另一个是路由器集群结构。

1) 高可扩展路由器

随着接入节点数的密集,汇聚或接入路由器的端口数量大量增加。当路由器的各种类型的线卡达到数百个的时候,一个机柜插入的线卡数已经不能满足,需要用多个机柜容纳大量的线卡。同时,由于网络流量的增加,对路由器提出 Tbps 到数十 Tbps 的交换需求,要求交换结构具有更高的扩展能力,可采用如图 5-54 所示的多机柜互连的高可扩展路由器结构。

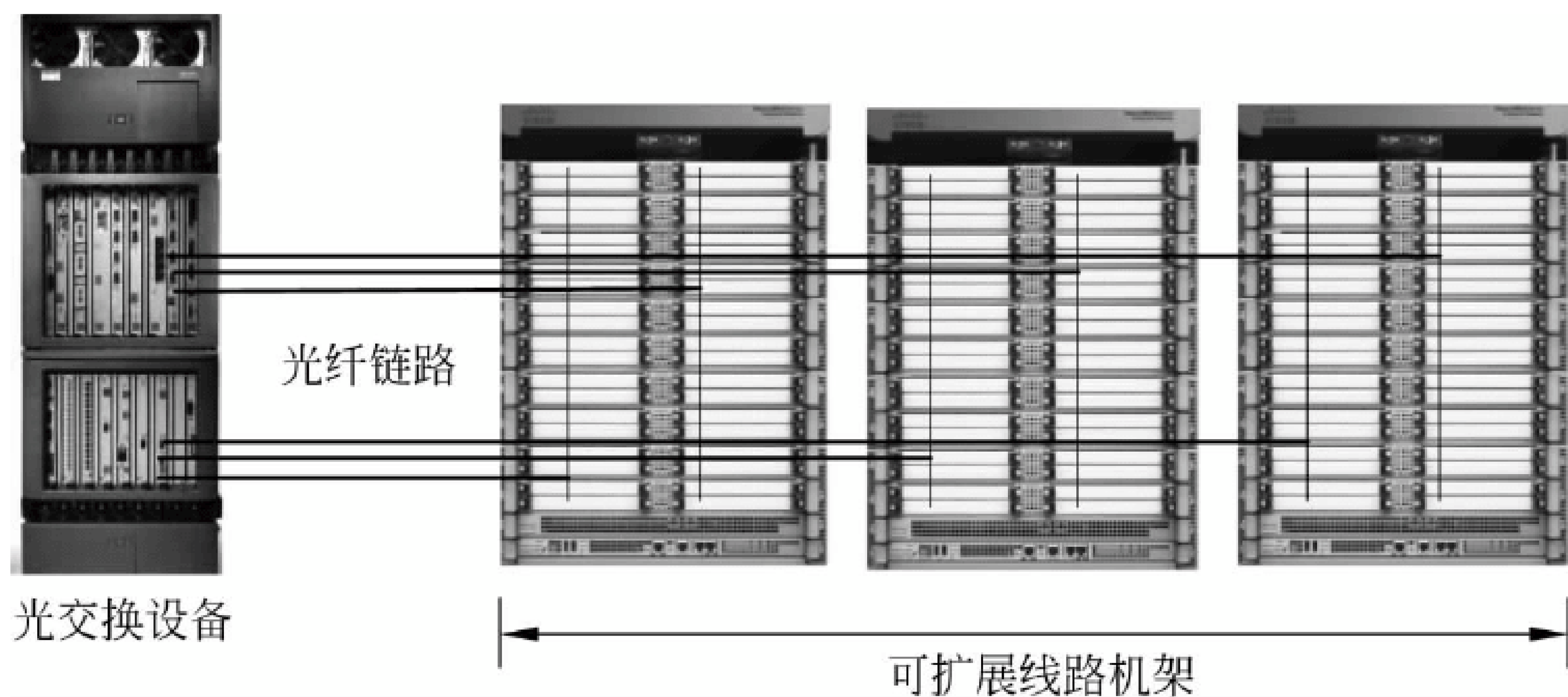


图 5-54 多机柜互连的高可扩展路由器结构示意图

2) 路由器集群结构

网络运营商在网络通信枢纽点(Point of Presence,POP)机房需要集中安放多台路由器,而集群技术是最有效的解决路由器扩展性问题的技术。路由器集群结构可以在不增加网络复杂性、方便维护的前提下,用比较少的投入来满足业务高速增长、网络性能及容量快速提升的需求。理解路由器集群结构需要注意以下几个问题。

第一,路由器集群技术是将两台或两台以上的核心路由器互连起来,对外只表现为一台逻辑路由器;内部采用相应的结构与并行算法,形成多级、多平面的交换矩阵系统,使得核心路由器之间能够协同工作与路由转发任务的并行处理。路由器集群结构又称为"路由器矩阵"技术或"多机箱(Multi-Chassis)组合"技术。

第二,根据集群设备数量的不同,路由器集群的结构可以分为两框集群、四框集群和多框集群(如图 5-55 所示)。

从实现技术的角度看,集群内多台路由器之间的连接,可以采用对等互连或中心交换框



图 5-55　路由器集群的结构

连接的方式。集群的具体结构形态又可分为“背对背”“1 拖 m”“2 拖 m”，以及“n 拖 m”的结构。“背对背”是指将两台用户框直接互连，不需要通过中心交换框互连；“1 拖 m”是指将 m 台用户框通过一台中心交换框进行集群互连；“2 拖 m”是指将 m 台用户框通过两台中心交换框进行集群互连。集群技术未来将朝着“n 拖 m”的结构发展，集群路由器的数量可以达到 64 框以上，系统容量也可达到 100Tbps 以上。

目前两框集群技术已经具备商用能力，四框及多框集群也将逐步成熟。对于两框集群技术，又可采用“背对背”“1 拖 2”及“2 拖 2”等方式来实现(如图 5-56 所示)。

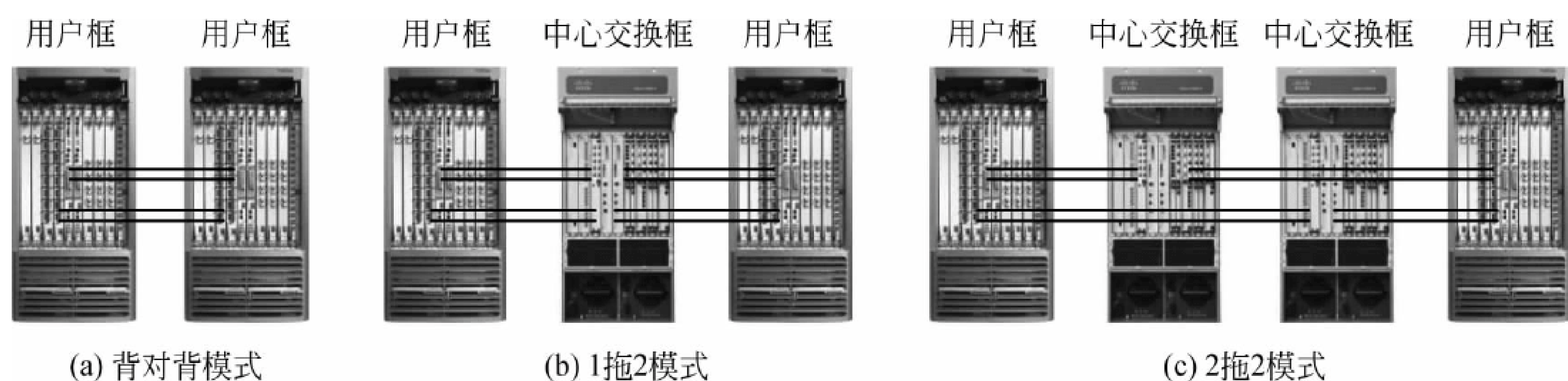


图 5-56　两框集群的三种实现方式

目前，常用的两框集群“背对背”结构如图 5-57 所示。很多校园网或宽带城域网都采用“背对背”结构。

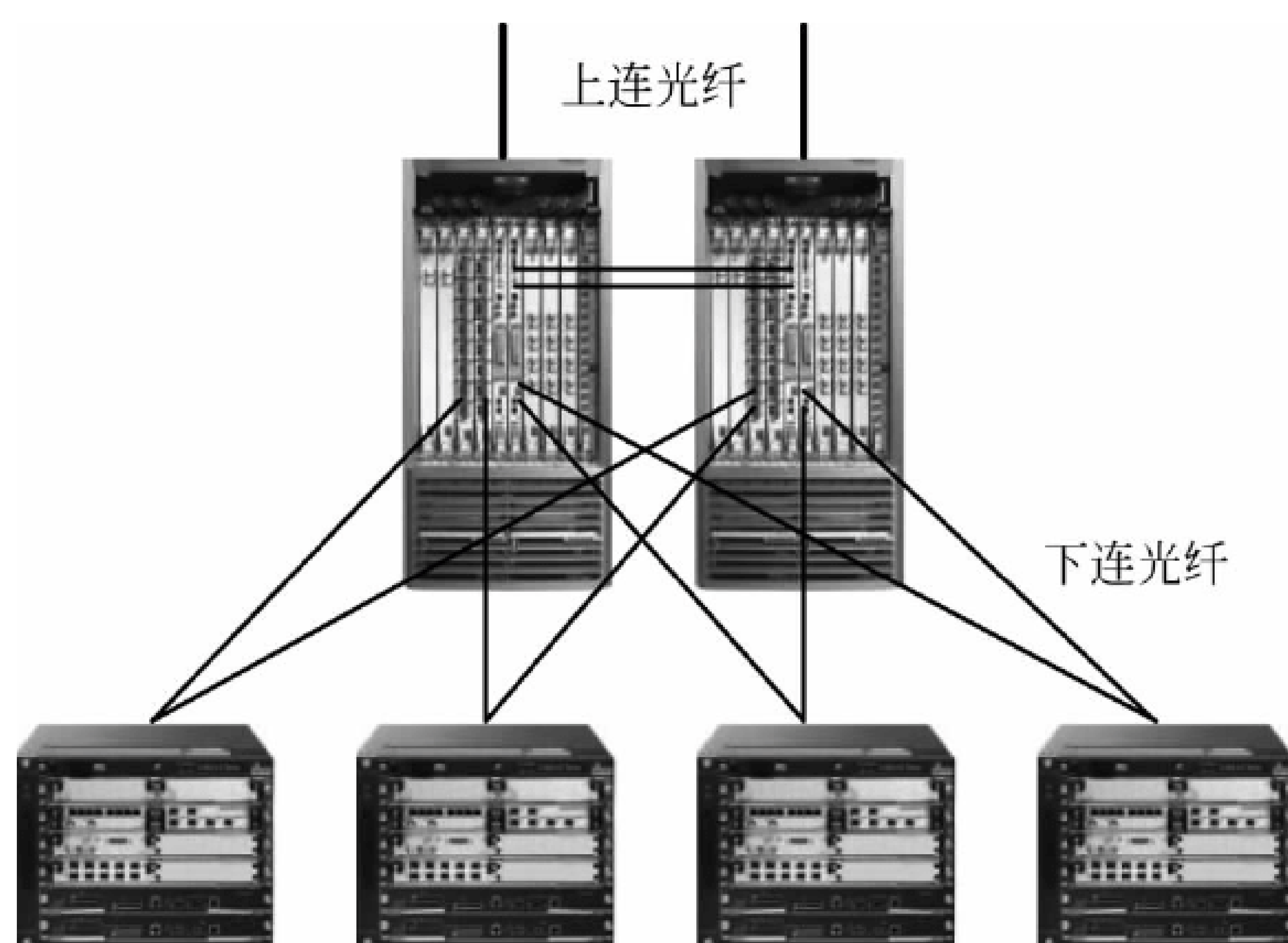


图 5-57　常用的“背对背”结构示意图

第三，由于集群系统中的每个路由器机箱之间都是通过高速光背板互连，因此需要采用专用的大容量数据光纤进行互连。目前用于路由器机箱之间连接的光缆，可以达到每一条光缆包含72芯2.5Gbps的光纤，光缆的传输能力为180Gbps。通过光纤实现路由器集群内部的互连，可以实现核心路由器容量平滑地扩充到原有的2倍、4倍、6倍、8倍甚至更高，而且不会增加路由的跳数和复杂度；核心路由器容量从几太比特每秒扩展到数十太比特每秒，很好地解决了核心层"大容量、可扩展"的问题。

第四，从网管的角度来看，路由器集群结构中有一个统一的管理和路由控制引擎，集群中的路由器属于一台逻辑路由器，从而使网络拓扑和路由结构变得简洁清晰，维护简单方便。

目前，核心路由器集群技术正逐步走向成熟，其商用能力、硬件结构的可靠性，以及软件体系的稳定性等方面，都还需要经受市场长时间的验证。

互联网、移动互联网与物联网的广泛应用与网络规模的不断扩大对路由器不断提出更高的要求，而路由器性能的改善只能从硬件与软件两个方面着手。在路由器硬件改进到一定程度时，继续提高路由器的性能、增加路由器的功能，只能从软件角度去考虑。目前，进一步拓展路由器性能与功能的研究重点在三个方面：路由器软件的可编程、虚拟化与可重用。

第三部分　习题参考答案

1. ① Switch　② Hub　③ Repeater　④ Bridge　⑤ Router

2. ① 05-2A-00-12-88-11
   ② 08-00-00-55-00-20
   ③ 09-2A-00-00-22-10
   ④ 0A-20-00-02-08-60
   ⑤ 128.4.0.2
   ⑥ 02-2B-01-50-26-66

3. 直接广播地址：193.0.255.255
   受限广播地址：255.255.255.255
   这个网络上的特定主机地址：0.0.5.1
   回送地址：127.0.0.0

4. ① RIP
   ② BGP-4
   ③ OSPF
   ④ BGP-4
   ⑤ OSPF

5. ① /12
   ② 125.144.0.0
   ③ 0.1.131.9
   ④ 125.159.255.255
   ⑤ 125.144.0.1

⑥ 125.144.255.254

6. （1）第一个子网的网络地址为：192.12.66.128，网络前缀为/26，可用网络地址为192.12.66.129～12.12.66.190，可用子网地址数大于50个。

（2）第二个子网的网络地址为：192.12.66.192，网络前缀为/27，可用网络地址为192.12.66.193～12.12.66.222，可用子网地址数大于20个。

（3）第三个子网的网络地址为：192.12.66.224，网络前缀为/27，可用网络地址为192.12.66.225～12.12.66.254，可用子网地址数大于20个。

7. 路由2：目的网络为195.192.0.0/17。

8. R1路由表如表5-13所示。

表5-13　R1路由表

掩　码	目的地址	下一跳地址	输出端口
255.255.255.0	200.8.4.0	—	m2
255.255.0.0	86.4.5.0	211.4.10.3	m1
255.255.0.0	129.4.6.0	211.4.10.3	m1
255.255.0.0	86.4.5.0	200.8.4.12	m2
255.255.0.0	129.4.6.0	200.8.4.12	m2
0.0.0.0	0.0.0.0	221.5.1.3	m0

9. 正确的是D选项202.168.2.0，255.255.255.0，202.168.1.2

10. （1）主机A与主机B属于同一个子网，它们可以不通过路由器直接通信。

（2）默认网关IP地址设置错误，导致主机A不能够与DNS服务器通信。解决的办法：将默认网关地址改为202.111.222.161或同一个子网的其他地址即可。

11. 分组头在传输过程中没有出错。

第 6 章 传输层

第一部分　学习目的、要求与知识点结构

1. 学习目的

计算机网络本质的活动是实现分布在不同地理位置的主机之间的进程通信，以实现应用层的各种网络服务功能。传输层的主要作用就是要实现分布式进程通信，因此它是整个协议结构的核心。本章将从分布式进程通信的基本概念出发，讨论传输层的基本功能，传输层向应用层提供的服务，以及实现这些服务的传输层协议 TCP 与 UDP 的基本内容。通过本章的学习，读者将掌握用户数据报协议 UDP 与传输控制协议 TCP 的基本内容，为读者进一步研究应用层与应用层协议打下基础。

2. 学习要求

（1）理解：网络环境中分布式进程通信的基本概念。

（2）掌握：进程通信中客户/服务器模式的基本概念。

（3）掌握：传输层的基本功能与服务质量 QoS 的基本概念。

（4）掌握：UDP 的基本内容。

（5）掌握：TCP 的基本内容。

3. 本章知识点的组织与结构

本章知识点的组织与结构如图 6-1 所示。

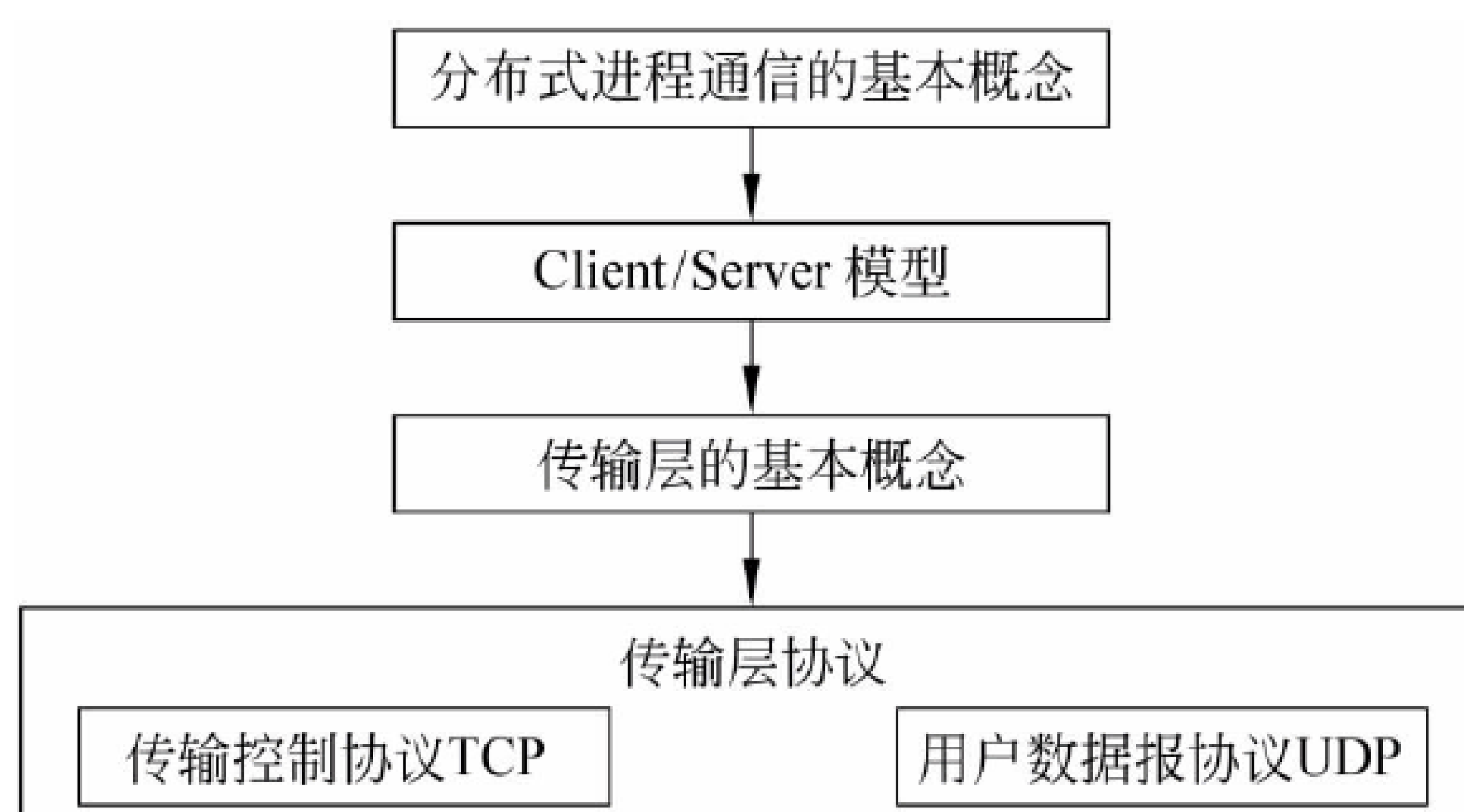


图 6-1　第 6 章知识点结构示意图

第二部分　教学内容问答

问题 6-1：为什么说传输层“端-端”通信是一次质的飞跃？

计算机网络的设计者在设计传输层时，希望能够达到以下目的。

(1) 网络层的 IP 地址标识了主机、路由器的位置信息；路由选择算法可以在 Internet 中选择一条源主机-路由器、路由器-路由器、路由器-目的主机的多段“点-点”链路组成的传输路径；IP 协议通过这条传输路径完成 IP 分组数据的传输。传输层协议是利用网络层所提供的服务，在源主机的应用进程与目的主机的应用进程之间建立“端-端”连接，实现分布式进程通信。

(2) Internet 中的路由器与通信线路构成了传输网(或承载网)。传输网一般是由电信公司运营和管理的。如果传输网提供的服务不可靠(例如频繁丢失分组)，用户无法对传输网加以控制。解决这个问题需要从两个方面入手：一是电信公司进一步提供传输网的服务质量；二是传输层对分组丢失、线路故障进行检测，并采取相应的差错控制措施，以满足分布式进程通信对服务质量(QoS)的要求。因此，传输层要改善 QoS，以达到计算机进程通信所要求的服务质量问题。

(3) 传输层可以屏蔽传输网实现技术的差异性，弥补网络层所提供服务的不足，使得应用层在设计各种网络应用系统时，只需要考虑选择怎样的传输层协议可以满足应用进程通信的要求，而不需要考虑数据传输的细节问题。

因此，从网络层的“点-点”通信到传输层的“端-端”通信是一次质的飞跃，为此传输层需要引入很多新的概念和机制。

问题 6-2：网络环境中分布式进程通信有哪些重要的特点？

要了解网络环境中的进程通信，需要与单机环境的进程通信做一个比较。

1. 单机系统中的进程通信方法

进程和进程通信是操作系统中的一个最基本的概念。在单机系统中，多个进程共享单一的 CPU，因此在一个时刻某个进程在使用 CPU，而有的进程在等待分配 CPU，而有的进程在等待其他的条件。一个进程在不同的时刻处于不同的状态。正在运行的进程叫作运行态；等待分配 CPU 的进程叫作就绪态，而等待其他条件的进程叫作等待态。进程的状态反映出进程执行过程的变化。要保证系统正常地工作，操作系统必须对进程的创建、撤销与状态转换进行控制。在操作系统的支持下，计算机系统中的各个进程互相独立地并发运行，但是它们要共享计算机资源，因此进程在运行过程中互相之间存在着互斥和同步的关系。

从进程的观点来看，操作系统是由可以同时独立运行的程序和一个对这些程序进行协调的核心所组成，这些同时运行的程序叫作进程，每一个进程都完成一种特定的任务，而操作系统的核心则是控制和协调这些进程的运行，解决进程之间的通信。

2. 网络环境中进程通信需要解决的主要问题

如果一台计算机连接到网络环境中，那么在网络环境中的计算机系统之间的进程通信与单机状态下的进程通信将有较大的差别。用一句最简单的话去描述计算机网络，那就是：计算机网络是分布在不同地理位置的多台独立的计算机系统的集合。“独立的计算机系统”意味着联网的每一台计算机的操作与资源是由自己的操作系统管理。用户共享的网络资源

及网络所能提供的服务功能最终是通过网络环境中的分布式进程通信来实现的。这种网络环境中的进程通信与单机系统内部的进程通信的主要区别在于网络中主机的高度自主性。因为网络中不同的主机系统之上,没有一个高层的操作系统进行统一的进程与资源的控制与管理,因此网络中一台主机对某一台主机的活动状态、位于这台主机系统中的各个进程状态、这些进程什么时间参与网络活动,以及它们希望与网络中哪一台主机的什么进程通信等情况一概无从知道。如果要实现网络环境中分布在不同主机系统中的进程间实现通信,必须对这些重要的问题提出解决的办法。

网络环境中分布式进程通信的实现必须解决以下三个主要问题:进程命名与寻址方法、多重协议的识别与进程间相互作用的模式。

问题6-3:如何解决网络环境中的进程标识问题?

为了回答这个问题,需要注意以下几点。

(1)网络环境中的进程通信要解决的第一个问题是进程标识。在一台计算机中,不同的进程可以用进程号或进程标识(Process ID)唯一地标识出来。网络环境中的进程标识需要使用主机地址。那么,网络环境中一个完整的进程标识应该是:本地主机地址-本地进程标识、远程主机地址-远程进程标识。

(2)从用户使用的角度来看,用户服务程序是用名字表述的,例如“文件传输服务FTP”“域名服务DNS”“电子邮件E-mail”等。但是计算机系统需要使用它可以理解的数字代码,即进程地址。因此,必须建立起“进程名字”与“进程地址”之间的映射关系,并且通过名字服务程序来完成进程名字与进程地址之间的转换。

(3)进程地址也叫作端口号(Port Number)。端口号是TCP与UDP与应用程序连接的访问点,是TCP与UDP软件的一部分。TCP/IP的传输层协议规定了一些标准的保留端口号,用于服务器进程;用户可以申请使用非保留端口,这些非保留端口的端口号在本机中也是唯一的。因此,端口号可以作为网络环境中的进程标识。

(4)多重协议的识别。

如果网络环境中的两台主机要实现进程通信,那么它们首先要约定好传输层协议类型。因此,考虑到进程标识和多重协议的识别,网络环境中一个进程的全网唯一的标识需要一个三元组来表示。这个三元组是(协议,本地地址,本地端口号)。在UNIX操作系统中,这个三元组又叫作半相关。网络环境中的进程通信需要涉及两个不同主机的进程,因此一个完整的进程通信标识需要由一个五元组来表示。这个五元组是(协议,本地地址,本地端口号,远地地址,远地端口号)。在UNIX操作系统中,这个五元组叫作一个相关。

问题6-4:如何理解进程间相互作用的Client/Server模式?

为了回答这个问题,需要注意以下几点。

1. 什么是Client/Server模式

(1)在计算机网络中,每台联网的计算机既要为本地用户提供服务,也要为网络的其他主机的用户提供服务。

网络的每项服务都对应一个“服务程序”进程。这些进程要为每个获准的网络用户请求执行一组规定的动作,以满足用户网络资源共享的需要。

(2)请求服务、发起本次进程通信的本地计算机进程叫作客户进程(Client),远程计算

机提供服务的进程叫作服务器进程(Server)。

(3) 网络环境中的进程通信采用Client进程发出服务请求,远程Server进程响应客户端请求并提供服务进程之间的通信模式,叫作Client/Server模式。

2. 采用Client/Server模式的理由

在TCP/IP体系中,进程间的相互作用采用Client/Server模式的理由主要有以下几点。

1) 网络资源分布的不均匀性

网络资源分布的不均匀性表现在硬件、软件和数据等三个方面。

(1) 硬件。

网络中主机系统类型、作用和能力存在着很大的差异。它可以是一台大型计算机、高档服务器,也可以是一台个人计算机,甚至是一个PDA或者是一个家用电器。它们在运算能力、存储能力和外部设备的配备等方面存在着非常大的差异。早期的无盘工作站本身没有硬盘,它每次启动时首先要访问一台主机,从这台主机下载启动程序,它在工作过程中产生的数据也必须保存在主机的硬盘中。

(2) 软件。

从软件的角度看,出于所属权、管理与运行环境要求等原因,很多大型应用软件都是安装在某台主机系统中,网络用户可以通过网络去访问,成为合法用户,然后提出和完成计算任务。

(3) 信息资源。

在网络中,某些信息以数据库方式集中存放在一台或几台具有收集、维护和更新特权的主机中,其他合法用户可以访问这些信息资源。这样做对保证信息使用的合法性、安全性,以及保证数据的完整性与一致性是非常必要的。

(4) 网络资源分布不均匀性是客观存在的。

网络资源分布的不均匀性是网络应用系统设计者的设计思想的体现。网络组建的目的就是要实现资源的共享,"资源共享"表现出网络中不同结点之间在硬件配置、运算能力、存储能力,以及数据分布等方面存在着差距与不均匀性。能力强、资源丰富的充当Server,能力弱或需要某种资源的成为Client。

因此,从网络资源分布的不均匀性角度来看,采用Client/Server模式是恰当的。

2) 网络环境中进程通信的异步性

(1) 网络环境中进程通信是异步性的。对于分布在不同主机系统中的进程,进程什么时间发出通信请求,希望和哪台主机的哪个进程通信,以及对方进程是否能接受通信请求,这些全然不知。

(2) 不存在一个统一调度与协调的高层操作系统。

(3) Client/Server模式的工作实质是"请求驱动"。每次通信由Client进程随机发起。

(4) Server进程从开机之时起就处于等待状态,以保证及时响应Client的服务请求。为了实现Server的功能,在Server设计中要解决Server的并发请求处理能力,并发Server的进程标识,以及Server安全等几个主要问题。

因此,从进程通信过程中的数据交换角度来看,采用Client/Server模式也是恰当的。

问题 6-5：如何实现进程通信中的 Client/Server 模式？

为了回答这个问题，需要注意以下几点。

1．Server 对并发请求的处理能力

（1）在网络环境中，Client 进程发出请求完全是随机的。在同一个时刻，可能有多个 Client 进程向一个 Server 发出服务请求。因此，Server 必须要有处理并发请求的能力。

（2）解决 Server 处理并发请求的方案基本上有以下两种：一是采用并发 Server 的方法；二是采用重复 Server 的方法。

（3）并发 Server 的核心是使用一个守护程序（Daemon），该程序在系统启动的时候随之启动。在没有 Client 的服务请求到达时，并发 Server 处于等待状态。当 Client 的服务请求到达时，Server 根据 Client 的服务请求的进程号，激活相应的子进程，由子进程为 Client 提供服务，而 Server 回到等待状态。

（4）Server 必须拥有一个全网熟知的进程地址。网络中的 Client 进程可以根据 Server 进程的熟知地址，向 Server 提出服务请求。在实现进程通信的过程中，Client 与 Server 进程分别形成自己的半相关的三元组，然后 Client 根据 Server 进程的熟知进程地址建立全相关的五元组。

（5）重复 Server 是通过设置一个请求队列来存储 Client 的服务请求。Server 采用先来先服务的原则，顺序处理 Client 的服务请求。

（6）并发 Server 适应于面向连接的服务类型，而重复 Server 适应于无连接的服务类型。

（7）由于 Server 的特殊地位，Server 控制着网络共享的资源，具有更高的权限，它要完成用户合法身份的识别、资源访问的管理，因此 Server 的安全性也就显得格外重要。它是系统安全性设计的一个重点问题。

2．UNIX 进程通信实现方法

我们可以以 BSD UNIX 为例，更直观地解释网络环境中进程通信的实现方法。

1）Socket 的基本概念

Socket 在很多软件书籍中被译成“套接字”“插口”与“接口”。在进程地址命名中本地的每个进程用一个半相关描述，即（协议，本地地址，本地端口）；一个完整的进程连接需要使用一个相关描述，即（协议，本地地址，本地端口，远地地址，远地端口）。在网络的讨论中，一个 Socket 定义为一个主机的 IP 地址与该主机中的一个进程的端口号。

在网络中，应用层可以利用 Socket 建立进程连接，实现数据交换。Socket 是面向 Client/Server 模式而设计的，针对 Client 和 Server 程序提供不同的 Socket 系统调用，Client 随机申请一个 Socket 号，Server 使用全局的熟知 Socket 号，Client 可以随时向 Server 发出服务请求。

2）UNIX Socket 调用

对于 UNIX 系统，Socket 调用与文件访问操作有很多的相似之处。文件访问是本地的输入/输出，文件号对应一个具体的块文件或字符文件。Socket 调用是网络的输入/输出。UNIX 文件访问操作常用的风格是：open-read-write-close，即打开文件、读/写文件与关闭文件。Socket 调用与它很类似。UNIX 主要的 Socket 调用如下。

(1) 创建 Socket——socket()。

应用程序在使用 Socket 之前必须首先拥有一个 Socket 号。这就相当于用户在安装电话时,首先要向电话局申请一个电话号码。

socket()向应用程序提供创建 Socket 的手段。socket()调用的格式是:

```
socketid= socket(af,type,protocol)
```

其中,返回值 socketid 是一个整数(即 Socket 号)。创建 Socket 实际上是申请一个属于自己此次进程通信的 Socket 号。socket()一共有以下三个参数。

① 地址族(address family,af)指出 socket 使用的地址类型。UNIX 支持的地址类型有:AF-UNIX,表示 UNIX 内部地址;AF-INET,表示 TCP/IP 地址;AF-NS,表示 Xerox NS 地址。

② 类型 type 是指创建 Socket 的应用程序所希望的通信服务类型。Socket 支持的通信服务类型有:sock-STREAM,表示流 Socket;sock-DGRAM,表示数据报 Socket;sock-RAW,表示原始 Socket;sock-SEQPACKET,表示定序分组 Socket。

③ 协议 protocol 是 Socket 请求使用的协议。地址与协议类型是相联系的。AF-INET 表示使用的是 TCP/IP 族。其中,sock-STREAM 表示流 Socket 使用 TCP,sock-DGRAM 表示数据报 Socket 使用 UDP。

(2) 指定本地地址——bind()。

调用创建 socket()是实现创建 Socket 通信的第一步,它只指定了相关五元组中的协议,而指定本地地址 bind()系统调用给出了本地地址与本地端口。

bind()系统调用的格式是:

```
bind(socketid,localaddr,addrelen)
```

其中,socketid 是本地的 Socket 号;localaddr 是本地地址,在 TCP/IP 族中它就是本地主机的 IP 地址;addrelen 对应于 IP 地址长度。

(3) 建立 socket 连接——connect()与 accept()。

connect()与 accept()调用完成两个相关的连接。其中,connect()用于建立连接。这里的连接有两个含义,一是指在两个 Socket 之间沟通,二是在传输层建立连接,如 TCP 连接。connect()调用主要为面向连接的传输服务设计,而 accept()调用完全是为面向连接的传输服务设计。无连接 Socket 进程可以调用连接 connect()。但是,此时本地系统与远地系统之间并没有真正建立连接。无连接 Socket 进程可以调用连接 connect(),实际上是通知操作系统,将来自指定 Socket 的数据送到本 Socket。

accept()调用用于面向连接的 Server,调用格式为:

```
newsock= accept(socketid,clientaddr,paddrlen)
```

其中,socketid 指本地 Socket 号;clientaddr 指向 Client 的地址;paddrlen 表示 clientsocket 长度;newsock 是 accept()调用后,返回的新的 Socket 号,它是指从 Server 的 Socket 号。

通过 socket()、bind()、connect()与 accept()调用,可以建立一个完整的相关五元组。socket()指定协议;bind()指定本地的地址与 Socket 号;connect()指定远程主机地址与远

程 Socket 号，accept()确认远程主机地址与远程 Socket 号。

(4) 接收 Socket 连接——listen()。

listen()调用是用于面向连接 Server，它表示同意接受连接。

listen()调用格式是：

```
listen(socketid,quelen)
```

其中，socketid 是本地 Socket 号，表示 Server 可以在此 Socket 号上接受服务请求；quelen 表示请求的队列长度。

(5) 发送数据——write()、writev()与 send()、sendto()、sendmsg()。

用于 socket 数据发送的系统调用有 5 个，其中，write()、writev()、send()用于面向连接的传输服务，sendto()、sendmsg()用于无连接的传输服务。

用于面向连接传输 write()、writev()、send()调用的格式比较一致，例如：

缓冲发送：write(socketid,buff,bufflen)

集中发送：writev(socketid,iovecor,vectorlen)

可控缓冲发送：send(socketid,buff,bufflen,flags)

其中，socketid 为本地 Socket 号；buff 与 bufflen 为发送缓冲区指针和发送缓冲区大小；iovecor 与 vectorlen 指向 I/O 向量表的指针与向量表的大小；flags 区别 write()与 send()。

(6) 接收数据——read()、readv()与 recvfrom()、recvmsg()。

接收数据调用 read()、readv()与 recvfrom()、recvmsg()和发送数据调用是一一对应的。不同之处在于，发送数据调用的 Buff 是指针，而接收数据调用中的 Buff 是实际读出的值。

问题 6-6：UDP 有哪些主要的特点？

UDP 的特点主要有以下几点。

1. UDP 是一种无连接的传输层协议

UDP 在完成进程之间的通信中，提供了有限的差错检验功能。设计一个比较简单的传输层 UDP 的目的，是希望以最小的开销来实现网络环境中的进程通信。

UDP 不提供差错纠正、队列管理、重复消除、流量控制和拥塞控制。它只通过校验和的方式提供差错检测。UDP 自身提供最小功能，应用程序要保证数据报文传输的准确性、按序传输必须自己解决。

传输层协议需要具有以下主要的功能：一是创建进程到进程的通信，UDP 使用端口号来完成这种通信；二是在传输层提供流控制机制，UDP 在一个非常低的水平上完成这个功能，在接收到分组时没有流量控制，也没有确认机制。UDP 是一种无连接的传输层协议，它只从进程接收数据单元，并将它们通过网络层交付给接收方。数据单元必须足够小，能够装进一个 UDP 分组中。如果一个进程打算发送一个很短的报文，同时它对该报文的可靠性要求不高，那么它就可以使用 UDP。RFC768 是 UDP 的正式规范，三十多年来没有做出过重大的修改。

2. UDP 与应用层协议的关系

TCP/IP 族的层次结构和层次之间有着严格的单向依赖关系，其内部关系是十分清晰

的。TCP/IP的应用层协议类型很多,但是它们之间存在着明确的依赖关系。例如,应用层使用UDP的主要有简单文件传送协议TFTP、远程过程调用RPC、网络时间协议NTP、引导协议BOOTP,域名服务DNS既可以使用UDP也可以使用TCP,而其他的应用层协议需要依赖TCP,这就是UDP和TCP端口号分配不一样的原因。UDP和TCP在网络层都是使用IP协议。图6-2给出了UDP与TCP/IP族的其他协议层次位置以及依赖关系。

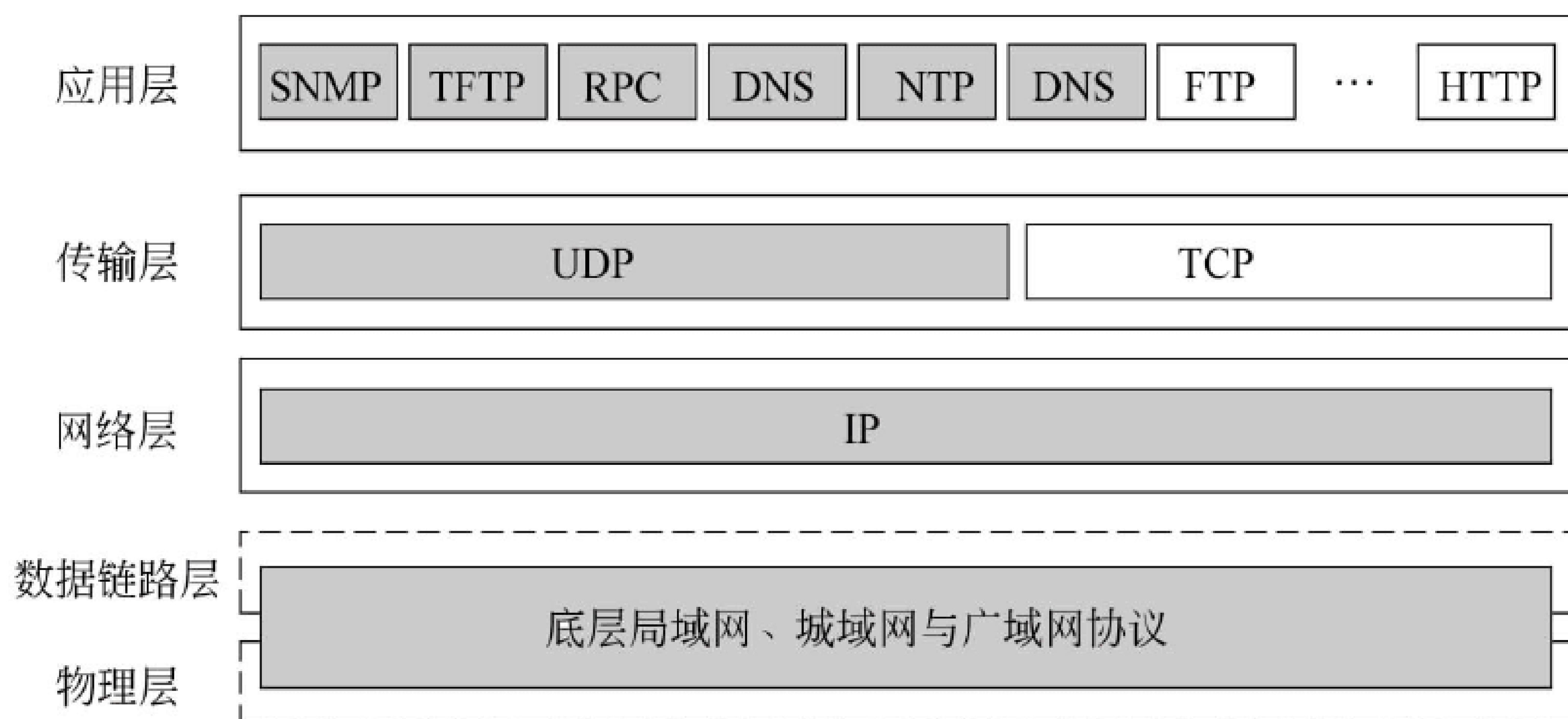


图6-2　UDP与其他协议的层次关系

问题6-7:如何理解UDP的基本工作过程?

理解UDP的基本工作过程,需要注意以下几点。

1. UDP用户数据报传输过程中的封装与拆封

(1) UDP提供无连接的服务,用户数据报在发送之前不需要有连接建立和释放的过程。这就表示UDP发送出的每个用户数据报都是一个独立的数据报,即使它们都来自相同的源进程,并且要发送到相同的目的进程,每个用户数据报都是独立的,用户数据报不进行编号,每个用户数据报可以走不同的路径。无连接服务的结果是,使用UDP的进程不能发送数据流,也不能期望UDP将这个数据流分割成为许多相关联的用户数据报。这就要求每个传输数据长度必须足够小,使它能装入到一个用户数据报中。因此,只有那些发送短报文的进程才应当使用UDP。

(2) UDP是一个不可靠的传输层协议,它没有流控制,因而也没有窗口机制。当到来的报文太多时,接收端可能会溢出。除检验和之外,UDP也没有差错控制机制,这表示发送端并不知道报文是丢失还是重复交付。当接收端使用检验和检测出错时,就将此用户数据报丢掉。缺少流控制和差错控制就表示使用UDP的进程的高层自己必须提供这些机制。

(3) 为了从一个进程将报文发送到另一个进程,UDP就要将报文进行封装和拆封。当进程有报文要通过UDP发送时,它就将此报文连同本地IP地址、本地端口号、远地IP地址、远地端口号以及数据长度传递给UDP。UDP收到数据后就加上UDP报头。UDP将这个用户数据报连同本地IP地址、本地端口号、远地IP地址、远地端口号一起传递给网络层IP软件。IP软件加上自己的IP报头,在协议字段使用值17,指出这个数据是从UDP来。这个IP数据报再传递给数据链路层。数据链路层收到IP数据报后,加上自己的帧头,再传递给物理层。物理层将这些比特编码为电信号或光信号,将其发送到远程的机器上。

2. UDP报文传输队列

1) Client端队列

UDP报文传输队列是与端口相关联在一起的。在Client端,当进程启动时,UDP将从

操作系统请求一个端口号,同时应创建一个出队列和一个输入队列。由Client创建的队列由被分配的临时端口号来标识,只要进程在运行,这些队列就起作用。当进程终止时,队列就被撤销。

Client进程使用在请求中指明的源端口号将报文发送到出队列。UDP逐个地将报文取出,加上UDP报头,交付给网络层。输出队列可能出现溢出。如果发生溢出,操作系统就要求Client进程在继续发送报文之前要等待。当报文到达Client时,UDP要检查对应于该用户数据报中目的端口号输入队列是否被创建。如果已经创建,UDP就将收到的用户数据报放在该队列的末尾。如果没有这样的队列,UDP就丢弃该用户数据报,并请求ICMP向Server端发送不可达报文。所有发送给一个特定Client的入报文,不管是来自相同的或不同的Server,都被放入同一个队列。输入队列也可能会溢出。如果发生溢出,UDP就丢弃这个用户数据报,并请求向Server发送端口不可达报文。

2) Server端队列

Server端创建队列的机制是不同的。

(1) server在开始运行时,使用它的熟知端口去创建输入队列和输出队列。只要Server进程在运行,这些队列就一直是打开的。

(2) 当报文到达Server进程时,UDP要检查对应于该用户数据报的目的端口号所对应的输入队列是否被创建。如果有这样的队列,UDP就将收到的用户数据报放在该队列的末尾。所有发送给特定Server程序的输入报文,不管是来自同样的或不同的Client进程,都被放入同一个队列。

(3) 如果没有这样的队列,UDP就丢弃该用户数据报,并请求ICMP向客户端发送端口不可达报文。输入队列可能会溢出。如果发生溢出,UDP就丢弃这个用户数据报,并请求向客户发送端口不可达报文。

(4) 当Server需要回答Client时,它就使用在请求中指明的源端口号将报文发送到输出队列。UDP逐个地将报文取出,加上UDP报头,然后交付给网络层。输出队列也可能出现溢出。如果发生溢出,操作系统就要求Server进程在继续发送报文之前要等待。

3) UDP的复用和分用

(1) 当主机运行TCP/IP族时,可能有多个进程想要使用UDP的服务。为了处理这种情况,UDP可以进行复用和分用。

(2) 在发送端,可能有多个进程需要发送用户数据报。UDP可以从不同的进程接收报文,这些进程是由分配给它们的端口号来区分的。在加上UDP报头之后,UDP将用户数据报送往网络层。

(3) 在接收端可以有多个进程接收用户数据报。这就是一对多的关系,因而需要分用。UDP从网络层接收用户数据报,经过差错检查、除去UDP报头之后,UDP根据端口号将每一个报文交付到适当的进程。

问题6-8:UDP端口号是如何分配的?

要理解这个问题,需要注意以下几点。

1. Client与Server进程的确定

进程通信的首要问题是解决进程标识方法。TCP/IP族中用端口号来标识进程。UDP在完成进程到进程的通信中采用的是Client/Server工作模式。发起本次进程通信、请求服

务的本地计算机进程叫作 Client 进程,远程计算机提供服务的进程叫作 Server 进程。

2. 端口号的分类

1) 端口号的数值范围

在 TCP/IP 族中,端口号是在 0～65 535 之间的整数。

2) 端口号的类型

Internet 赋号管理局(IANA)定义的 UDP 端口号分成三类,即:临时端口号、熟知端口号和注册端口号。

(1) 临时端口号。

Client 程序定义它自己使用的端口号,这是由运行在 Client 主机上的 UDP 软件随机选取的,我们将它叫作临时端口号。临时端口号值的范围为 49 152～65 535,它们可以由任何进程来使用。

(2) 熟知端口号。

Server 进程也必须用一个端口号来定义。这个端口号不能随机选取。TCP/IP 给每种 Server 程序分配了确定的全局端口号,我们把它叫作熟知端口号或公认端口号。每个 Client 进程都知道相应的 Server 进程的熟知端口号。熟知端口号值的范围为 0～1023,它是统一分配和控制的。

UDP 的熟知端口号如表 6-1 所示。UDP 服务与端口号的映射表定期在 RFC768 等文本中公布,并可以在大多数 UNIX 主机的/etc/services 文件中得到。

表 6-1　UDP 的熟知端口号

端口号	服务进程	说明
7	Echo	将收到的数据报回送到发送器
9	Discard	丢弃任何收到的数据报
11	Users	活跃的用户
13	Daytime	返回日期和时间
17	Quote	返回日期的引用
19	Chargen	返回字符串
53	Nameserver	域名服务
67	Bootps	下载引导程序信息的 Server 端口
68	Bootpc	下载引导程序信息的 Client 端口
69	TFTP	简单文件传送协议
111	RPC	远程过程调用
123	NTP	网络时间协议
161	SNMP	简单网络管理协议
162	SNMP	简单网络管理协议

(3) 注册端口号。

注册端口号值的范围为 1024～49 151,用户根据需要可以在 IANA 注册,以防止重复。同时需要注意的是:TCP/IP 之外的其他操作系统可能使用与 IANA 不一样的熟知端口号和临时端口号。

问题 6-9:UDP 做检验和时为什么要加上伪报头?

要解释这个问题,需要注意以下几点。

（1）UDP 检验和字段是可选项，是用来检验整个用户数据报（包括报头）在传输中是否出现差错。这一点正反映出设计者效率优先的思想。因为计算检验和肯定是要花费时间的，如果应用进程对通信效率的要求高于可靠性时，应用进程可以不选择检验和。

（2）UDP 检验和包括三个部分：伪报头、UDP 报头以及应用层数据。伪报头是 IP 分组报头的一部分，其中填充域字段要填入 0，目的是使伪报头的长度为 16b 的整数倍。协议号域含有协议类型码，协议号 17 表示 UDP。UDP 长度含有 UDP 数据报的长度，不包括伪报头的长度。

（3）使用伪头部是为了验证 UDP 数据报是否传到正确的目的进程。UDP 数据报目的方的地址应该包括两部分：目的主机 IP 地址和目的端口号。UPP 数据报本身只包含目的端口号，由伪头部补充目的主机 IP 地址部分。UDP 数据报发送、接收端计算校验和时均加上伪头部信息。假如接收端发现校验和正确，则在一定程度上说明 UDP 数据报到达了正确主机上的正确端口。UDP 伪报头来自于 IP 报头，因此在计算 UDP 校验和之前，UDP 首先必须从 IP 层获取有关信息。这说明 UDP 与 IP 之间存在一定程度的交互作用。在 UDP/IP 这个协议结构中，UDP 校验和是保证数据正确性的唯一手段。

（4）计算检验和时，在 UDP 用户数据报之前要增加 12B 的伪头部。所谓伪头部是因为它本身并不是在 UDP 用户数据报的真正头部，只是在计算时，临时和 UDP 用户数据报连接在一起的。伪头部只在计算时起作用，它既不向低层传输，也不向高层传送。这一点正反映出设计者效率优先的思想。因为计算检验和肯定是要花费时间的，如果应用进程对通信效率的要求高于可靠性时，应用进程可以不选择检验和。

问题 6-10：UDP 适应于哪些应用领域？

UDP 校验和检错能力不是很强，但是算法简洁，运算速度快。UDP 适合于以下的应用领域。

1. 视频播放应用

在 Internet 上播放视频，用户最关注的是视频流能够尽快和不间断地播放，丢失个别数据报文对视频节目的播放效果不会产生重要的影响。如果采用 TCP，它可能因为重传个别丢失的报文而加大传输延迟，反而会对视频播放造成不利的影响。因此，对于视频播放程序这种对数据实时交付的要求高于可靠性的应用，UDP 更为适用。

2. 简短的交互式应用

有一类应用只需要进行简单的请求与应答报文的交互，客户端发出一个简短的请求报文，服务器端回复一个简短的应答报文，在这种情况下应用程序应该选择 UDP。应用程序可以通过设置“定时器/重传机制”来处理由于 IP 数据分组丢失问题，而不需要选择有确认/重传的 TCP，以提高系统的工作效率。

3. 多播与广播应用

UDP 支持一对一、一对多与多对多的交互式通信，这点 TCP 是不支持的。UDP 头部长度只有 8B，比 TCP 头部长度 20B 短。同时，UDP 没有拥塞控制，在网络拥塞时不会要求源主机降低报文发送速率，而只会丢弃个别的报文。这对于 IP 电话、实时视频会议应用来说是适用的。由于这类应用要求源主机以恒定速率发送报文，在拥塞发生时允许丢弃部分报文。

当然，任何事情都有两面性。简洁、快速、高效是 UDP 的优点，但是由于它不能提供必

需的差错控制机制,同时在拥塞严重时缺乏必要的控制与调节机制。这些问题需要使用 UDP 的应用程序设计者在应用层设置必要的机制加以解决。UDP 是一种适用于实时语音与视频传输的传输层协议。

问题 6-11:为什么能够查到很多关于 TCP 的 RFC 文档?

事实的确如此,这也正说明:设计者希望将 TCP 修改成一种功能完善的传输层协议。从多种 TCP 版本,以及几十种对 TCP 的功能扩充调整的 RFC 文档,以及很多关于 TCP 的研究论文中可以看出,TCP 是最受人们关注的传输层协议。

多年来,除了对 TCP 进行的修改和功能扩充外,IETF 还公布了十几个与 TCP 有关的 RFC 文档。相对而言,UDP 的修改比较少,RFC 文档也比较简单。如果读者在站点 www.rfc-editor.org 查询与 TCP 文档有关的 RFC 的话,查询可以返回多达一百多个。对于网络课程学习的读者,可以提供有用的参考资料的 RFC 文档主要如下。

RFC793:TCP(协议标准)

RFC1144:用于低速串联链路的 TCP/IP 压缩报头(协议草案)

RFC1146:扩充的 TCP 校验和选项(协议草案)

RFC1263:TCP 功能扩充的问题(协议草案)

RFC1693:TCP 功能扩充的部分有序的服务(协议草案)

RFC1323:TCP 性能扩充(协议草案)

RFC2018:TCP 可选确认功能扩充(协议草案)

RFC2267:网络入口过滤,检测使用 IP 源地址欺骗算法的拒绝服务攻击(协议草案)

RFC2398:FYI0033TCP 程序测试工具(协议草案)

RFC2414:TCP 初始窗口的递增(协议草案)

RFC2415:递增初始 TCP 窗口长度的模拟实验(协议草案)

RFC2416:TCP 启动时把 4 个数据包放入三个缓冲区的实现方法(协议草案)

RFC2488:(BCP0028)借助标准机制使用卫星信道的 TCP 增强协议(协议草案)

RFC2525:TCP 程序设计问题(协议草案)

RFC2581:TCP 拥塞控制(协议草案)

RFC2582:TCP 快速恢复算法补充(协议草案)

RFC2757:传输延迟时间长的网络(协议草案)

RFC2760:卫星电路的 TCP 的研究(协议草案)

RFC2861:TCP 拥塞窗口的确认(协议草案)

RFC2883:TCP 的 SACK 选项功能扩充(协议草案)

RFC2923:TCP 与路由 MTU 发现(协议草案)

RFC2988:TCP 重传定时计算方法(协议草案)

RFC3042:借助有限传输提高 TCP 丢失恢复功能(协议草案)

RFC3155:有差错链路的端到端性能(协议草案)

对主要的 RFC 文档进行分析之后,可以看出,人们对 TCP 的扩充主要集中在改进 TCP 在网络拥塞下的恢复性能,减少因网络拥塞而引发的传输错误,以及如何选择传输窗口、接收窗口、超时数值、报文段长度等 TCP 变量的最佳值。另外的问题是在不同的通信子网中使用的 TCP 上。例如,在宽带网络与窄带网络中使用的 TCP,以及在移动通信信道与

卫星通信信道中使用的增强 TCP。在使用卫星通信信道时，又可以分为高轨道卫星与低轨道卫星的 TCP 问题。

TCP 是一个非常复杂的协议，与仅支持简单报文传输的 UDP 相比，TCP 可以提供人们想得到的几乎所有的传输层功能。可以预见，TCP 将在未来的 Internet 应用中继续发挥其不可替代的作用。值得注意的是，近年来，随着新的传输层协议 SCTP 的出现，越来越多的网络工程师开始考虑这种新的选择。

问题 6-12：TCP 具有哪些主要的特点？

尽管 TCP 和 UDP 都使用相同的网络层 IP 协议，但是 TCP 向应用层提供与 UDP 完全不同的服务。TCP 是一种面向连接的、可靠的传输层协议。TCP 在应用层和网络层之间，它在 IP 服务的基础上，增加了面向连接和可靠性的特点，提供面向连接的流传输。TCP 的特点主要表现在以下几个方面。

1. 支持面向连接的传输服务

如果将 UDP 提供的服务比作发送一封平信的话，那么 TCP 所能提供的服务相当于人们打电话。UDP 是一种可满足最低传输要求的传输层协议，而 TCP 则是一种功能完善的传输层协议。

面向连接对提高系统数据传输的可靠性是很重要的。应用程序在使用 TCP 传送数据之前，必须在源进程端口与目的进程端口之间建立一条 TCP 传输连接。每个 TCP 传输连接用双方端口号来标识；每个 TCP 连接为通信双方的一次进程通信提供服务。由于 TCP 建立在不可靠的网络层 IP 协议之上，IP 协议不能提供任何可靠性保障，因此 TCP 的可靠性需要自己来解决。

2. 支持字节流的传输

TCP 支持字节流传输。流(Stream)相当于一个管道，从一端放入什么内容，从另一端可以照原样取出什么内容。它描述了一个不出现丢失、重复和乱序的数据传输过程。如果用户是通过键盘输入数据，那么应用程序将逐个地将字符提交给发送端。如果数据是从文件得到，那么数据可能是逐行或逐块交付给发送端。应用程序和 TCP 每次交互的数据长度可能都不相同，但 TCP 是将应用程序提交的数据看成是一连串的、无结构的字节流。为了能够提供字节流方式的传输，发送端和接收端都需要使用缓存。发送端使用发送缓存存储从应用程序送来的数据。发送端不可能为发送的每个写操作创建一个报文段，而是选择将几个写操作组合成一个报文段，然后提交给 IP 协议，由 IP 协议封装成 IP 分组之后传输到接收端。接收端 IP 协议将接收的 IP 分组拆封之后，将数据字段提交给接收端 TCP。接收端 TCP 将接收的字节存储在接收缓存中，应用程序使用读操作将接收的数据从接收缓存中读出。

TCP 在传输过程中将应用程序提交的数据看成是一连串的、无结构的字节流，因此接收端应用程序数据字节的起始与终结位置必须由应用程序自己确定。

3. 支持全双工通信

TCP 允许通信双方的应用程序在任何时候都可以发送数据。由于通信的双方都设置有发送和接收缓冲区，应用程序将要发送的数据字节提交给发送缓冲区，数据字节的实际发送过程由 TCP 来控制；而接收端在正确接收到数据字节之后，将它存放到接收缓冲区，高层应用程序从缓冲区中读取数据。

4. 支持同时建立多个并发的TCP连接

TCP需要支持同时建立多个连接,这个特点在服务器端表现得更为突出。根据应用程序的需要,TCP支持一个服务器与多个客户端同时建立多个TCP连接,也支持一个客户端与多个服务器同时建立多个TCP连接。TCP软件将分别管理多个TCP连接。在理论上,TCP可以支持同时建立的上百,甚至上千条这样的连接,但是建立并发连接的数量越多,每条连接共享的资源就会越少。

5. 支持可靠的传输服务

TCP是一种可靠的传输服务协议,它使用确认机制检查数据是否安全和完整地到达,并且提供拥塞控制功能。TCP支持可靠数据传输的关键是对发送和接收的数据进行跟踪、确认与重传。需要注意的是:TCP建立在不可靠的网络层IP协议之上,一旦IP及以下层出现传输错误,TCP只能不断地进行重传,试图弥补传输中出现的问题。因此,传输层传输的可靠性是建立在网络层基础上,同时也就会受到它们的限制。

因此,总结以上讨论可以看出TCP的特点是:面向连接、面向字节流、支持全双工、支持并发连接、提供确认/重传与拥塞控制。

问题6-13:为什么TCP与UDP熟知端口号大多数是奇数?

TCP端口号分配方法与UDP原则上基本相同,只是根据应用层协议的关系,具体的应用类型是不同的。TCP端口号也是在0~65 535之间的整数。运行在远程计算机上的Server必须使用公认的熟知端口号。表6-2给出了TCP使用的一些熟知端口号。

表6-2 TCP常用的熟知端口号

端口号	服务进程	说明
7	Echo	将收到的数据报回送到发送器
9	Discard	丢弃任何收到的数据报
11	Users	活跃的用户
13	Daytime	返回日期和时间
17	Quote	返回日期的引用
19	Chargen	返回字符串
20	FTP	数据文件传送协议(数据连接)
21	FTP	控制文件传送协议(控制连接)
23	Telnet	虚拟终端网络
25	SMTP	简单邮件传送协议
53	DNS	域名Server
67	BOOTP	引导程序协议
80	HTTP	超文本传送协议
111	RPC	远程过程调用

细心的读者会发现:TCP与UDP的熟知端口号大多为奇数。这是有历史原因的。因为这些端口号都是从早期的网络控制协议NCP的端口号派生出来的。由于NCP不是全双工的,因此每种应用需要两个连接,并且两个连接使用奇偶成对的端口号。当TCP与UDP成为传输层的标准协议时,每个应用只需要一个端口号,原来的NCP使用的奇数端口号就被沿用下来,因此就出现TCP与UDP的熟知端口号大多为奇数的现象。

问题 6-14：TCP 在进程交互过程中使用了几种计时器？

为了实现 TCP 的功能，TCP 使用了 4 种计时器：重传计时器、坚持计时器、保持计时器和时间等待计时器。

1. 重传计时器

为了控制丢失的或丢弃的报文段，TCP 使用了重传计时器。重传计时器用来处理报文段的确认与等待重传的时间。当 TCP 发送报文段时，它就创建该特定报文段的重传计时器。在这之后可能会发生两种情况：如果在计时器截止时间到之前收到了对该报文段的确认，则撤销此计时器；如果在收到了对该报文段的确认之前计时器截止时间到，则重传该报文段并将计时器复位。

由于实际的两个 TCP 传输连接之间可能只相隔一个物理网络，也可能相隔了数千个互联的物理网络。因此一个传输连接所经历的路径长度，可以和另一个传输连接所经历的路径长度相差非常大，这就表明 TCP 不能对所有的连接使用相同的重传时间。对于两个 TCP 传输连接之间只相隔一个物理网络来说，重传时间也不可以是固定的。在通信量不太大的情况下，在一个连接上发送报文段和接收确认快。

累计确认的缺点是发送方不能获得关于所有成功的段传输的信息。假如前面尚有数据未得到确认，则后面的所有成功传输的段也得不到确认。例如，发送方发出两个报文段，第二个报文段传输成功，第一个报文段失败，在这种情况下，发送方将得不到任何确认，必须重传。在重传时采取什么策略是另一个问题：是两个报文段一起重传，还是逐段重传？两个报文段一起重传显然会造成浪费。逐段重传，传一段等待一个确认，再传下一段，又回到简单的停等协议方式。二者的效率都不高。

影响确认超时重传最关键的因素在于定时时间片的大小。Internet 环境中，要确定合适的定时时间片是一件相当困难的事情。一方面，Internet 进程通信既可能就在局域网上进行，也可能要穿越许多各种各样的中间网络，传输延迟变化范围相当大，另一方面，不同进程对之间的通信延迟还取决于不同信道的负载情况。总之，从发出数据到收到确认所需的往返时间(RTT)呈动态变化，很难把握。为适应上述情况，TCP 采用一种适应性重传算法。适应性重传算法的基本思想是：TCP 监视每一条连接的性能，由此推算出合适的时间片，当连接性能发生变化时，TCP 随即改变时间片值。

2. 坚持计时器

为了对付零窗口大小通知，TCP 需要另一个计时器。假定接收方的 TCP 宣布了窗口大小为零。发送方的 TCP 就停止传送报文段，直到接收方的 TCP 发送确认并宣布一个非零的窗口大小。这个确认可能会丢失。在 TCP 中，对确认报文段是不需要确认的。如果确认报文段丢失，接收方的 TCP 就认为它已经完成任务，并等待着发送方的 TCP 发送更多的报文段。发送方的 TCP 由于没有收到确认，就等待对方发送确认来通知窗口的大小。对方的 TCP 都在永远地等待着对方，这就可能出现了死锁。

要打开这个死锁，TCP 为每个连接使用一个坚持计时器。当发送方的 TCP 收到一个窗口大小为零的确认时，就需要启动坚持计时器。当坚持计时器期限到时，发送方的 TCP 就发送一个特殊的报文段，称为探测报文段。这个报文段只有 1 字节的数据。它有一个序号，但它的序号永远不需要确认，甚至在计算对其他数据确认时该序号也被忽略。探测报文段提醒接收方的 TCP：确认已丢失，必须重传。

坚持计时器的值设置为重传时间的数值。但是,若没有收到从接收方来的响应,则需发送另一个探测报文段,并将坚持计时器的值加倍和复位。发送方继续发送探测报文段,将坚持计时器的值加倍和复位,直到这个值增大到门限值(通常是60s)为止。在这以后,发送方每隔60s就发送一个探测报文段,直到窗口重新打开。

3. 保持计时器

保持计时器又叫作激活计时器,它是用来防止在两个TCP之间的连接长期空闲。假设Client建立了到Server的连接,传输了一些数据,然后就停止了传输,可能这个Client出故障了。在这种情况下,这个连接将永远地处于打开状态。为了解决这种问题,在大多数的实现中都是使Server设置激活计时器。每当Server收到Client的信息,就将计时器复位。超时通常设置为两小时。如果Server超过两小时还没有收到Client的信息,它就发送探测报文段。如果发送了10个探测报文段(每个相隔75s)还没有响应,就假定Client出了故障,因而就终止该连接。

4. 时间等待计时器

时间等待计时器是在连接终止期间使用的。当TCP关闭一个连接时,它并不认为这个连接马上就真正地关闭了。在时间等待期间中,连接还处于一种过渡状态。时间等待计时器的值通常设置为一个报文段的寿命期待值的两倍。

问题6-15:为什么TCP在计算校验和时要加上伪报头?

一般的协议设计时校验的目的是发现这一层数据传输单元在传输的过程中是否出现错误,而TCP则不同,它在计算校验和时在TCP报头之前加上了网络层的源IP地址、目的IP地址与协议TCP段长度等12个字节的伪报头。在发送端需要计算一次伪报头,接收端再计算一次伪报头,一致则表明传输过程中没有出错。这样做的目的有以下三点。

1. 防备不正确的报文段交付

如果接收的主机IP地址与伪报头中的目的地址不一致,则校验和计算结果一定与发送端计算的校验和不同,防备不正确的报文段交付。

2. 防备不正确的协议

如果因为某种原因,造成其他非TCP的报文数据错误地传送到TCP的主机,通过计算带有协议类型字段数据的伪报头,可以立即发现。

3. 防备不正确的段长度

如果TCP报文在传输过程中丢失了部分字节,就会使接收的TCP段长度出错,通过校验和计算也能够立即发现段长度不正确的问题。

对于伪报头的方法,也一直是有两种意见。一是从早期传输出错几率较高的角度,有人认为:TCP通过不额外传输伪报头,而是从IP分组头中复制重要字节内容的方法,用比较简单的方法,实现了必要的校验功能。另一种意见是:发现不正确的报文段交付问题应该有IP协议解决,设计伪报头实际上是不信任IP协议的可靠性,同时也违反了不隔层通信的原则,这样做是多此一举。

问题6-16:TCP默认的最大段长度MSS是多少?

要回答这个问题,需要注意以下几个问题。

1．TCP 最大段长度 MSS 的概念

应用层会话进程之间是通过 TCP 报文的数据字段以字节流的形式在传输数据的。由于 TCP 数据字段长度是可变的，那么就出现一个有趣的问题：我们究竟应该在每个报文段中放多少个字节的数据？

决定一个段发送多少数据的两个重要因素：一是接收端滑动窗口机制的当前状态；二是主机进程每次能够接收和处理的数据长度。TCP 需要结合这两个因素来确定每个报文段中应该放多少个字节的数据。无论当前窗口有多大，每个报文段中放入的字节不要超过这个值。因此人们将这个值称为最大长度(Maximum Segment Size，MSS)。最大长度的名称会产生误导。实际上，MSS 是 TCP 根据主机保持 TCP 段数据的存储空间，结合窗口状态选择的一个值，如果选择这个值是 100B，加上 TCP 报文头 20B，那么 TCP 报文段最大长度就是 120B；也可能下一次选择的值是 500B，那么 TCP 报文段最大长度就是 520B。

2．MSS 的默认值

从提高协议效率的角度，当然是希望段长度越多越好。同时也需要注意，如果 TCP 段长度选择大了，那么 TCP 报文段经过 IP 协议处理时就需要分片。为了不被分片，TCP 报文段最大长度需要考虑 IP 分组长度的限制。IPv4 分组的默认长度值是 576B，那么减去 IP 分组头的 20B、TCP 报文头的 20B，TCP 的默认 MSS 值为 536B。

问题 6-17：TCP 与 HDLC 协议都使用了确认与窗口机制，它们之间的区别是什么？

在讨论数据链路层的 HDLC 协议，以及传输层的 TCP 时都涉及保证传输可靠性与流量控制的确认与窗口机制，但是它们之间的区别主要表现在以下几个方面。

(1) HDLC 协议属于数据链路层协议。它的两个重要特点是：一是 HDLC 的活动窗口协议是实现相邻点-点链路之间的流量控制；二是它确认与窗口控制机制是针对帧的。

(2) TCP 属于传输层协议。它的两个重要特点是：一是 TCP 的活动窗口协议是实现源主机进程到目的主机进程之间端-端的流量控制；二是它确认与窗口控制机制是针对报文段的字节数。

显然，TCP 的确认与窗口机制要比链路层的 HDLC 协议复杂得多。

问题 6-18：实时传输协议 RTP/RTCP 的研究背景是什么？

为了增加任课教师对物联网技术的了解，教师用书补充了一部分物联网对传输层实时性要求高的实时传输层协议的内容。理解实时传输协议 RTP/RTCP 研究的背景，需要注意以下几个问题。

1．实时多媒体通信与非实时多媒体通信

多媒体数据类型分为语音、图形、图像与视频等。基于 Internet 的网络多媒体技术已经广泛应用于人类的工作、学习、通信、娱乐、科学研究、医疗与军事等各个领域。移动互联网的应用使得网络视频点播、网络音乐、网络电视、网络电影、网络广播、网络游戏、网络广告、网络地图发展到更高的阶段。

如图 6-3 所示，并非所有的多媒体网络应用都是实时的。如图 6-3(a)所示，我们经常用 HTTP 从 Web 服务器上下载视频节目。视频节目是首先录制好，存放在 Web 服务器上，我们下载到本地计算机上再去观看。这样的应用中，视频的录制与传输、观看不是发生在同一个时间，因此这种情况属于非实时的多媒体通信。在如图 6-3(b)所示的视频会议应用中，

摄像机连接在服务器上,它所传输的视频数据流以帧的形式,直接通过服务器与客户机的连接,传送到客户机上,在屏幕上显示出来。如果忽略数据流通过网络所产生的传播延时的话,那么视频数据流的产生、传输、接收、显示可以认为是发生在同一时间,这种情况属于实时的多媒体通信。

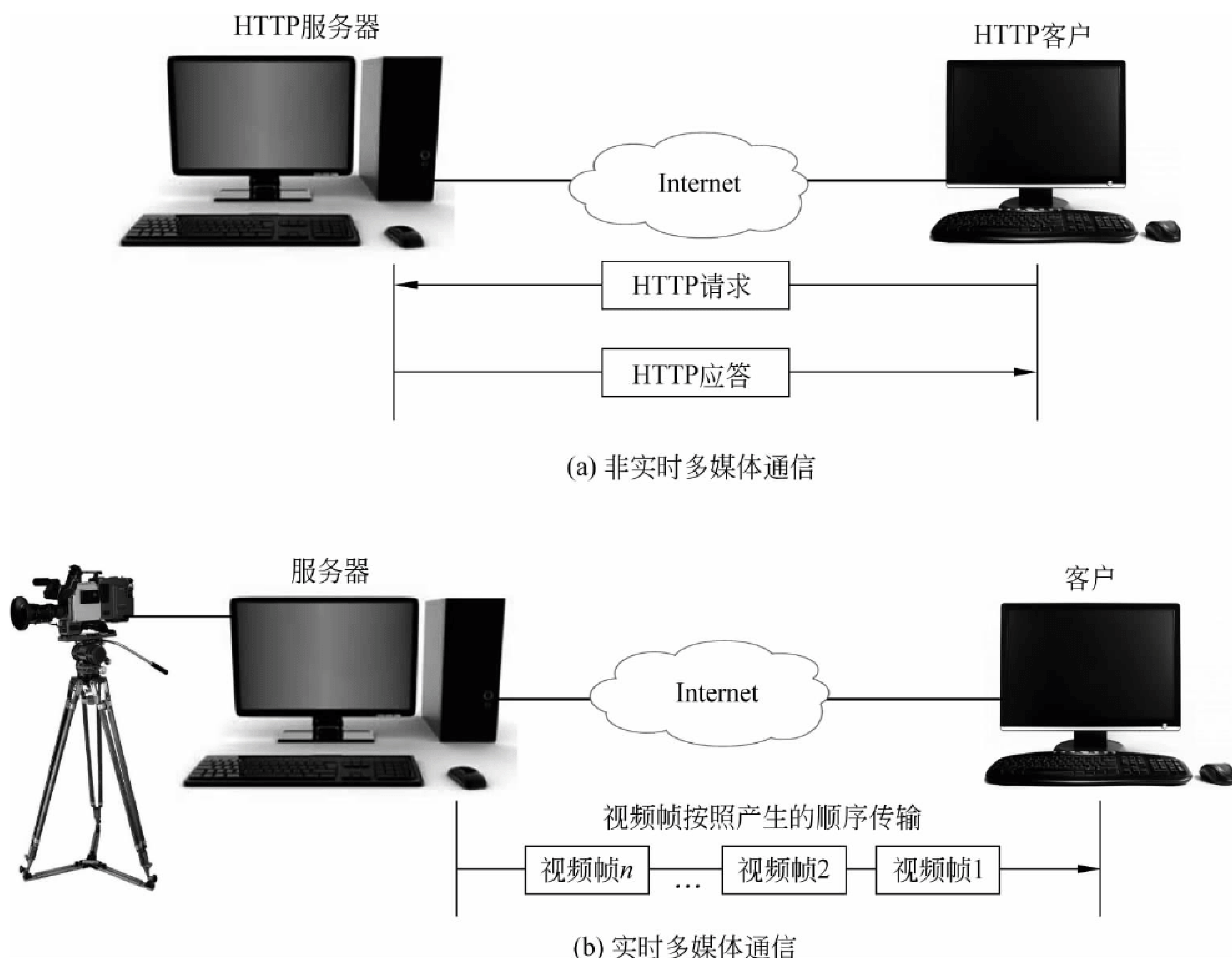


图 6-3 非实时与实时多媒体通信

2. 网络延时与延时抖动

网络多媒体应用与传统的网络应用对网络传输特性的要求有很大的不同。传统的网络应用,如 Web、E-mail、FTP、Telnet 应用对传输延时的要求相对不高,但是对数据传输的正确性要求很高。但是,在网络多媒体应用(如 IP 电视、IP 电话)中,对网络端-端延时与延时抖动就会很敏感,有些数据在延时之后就不能使用。延时与抖动如图 6-4 所示。在这个例子中,假定视频服务器产生的直播视频流封装在三个报文中。每个报文保留有 10s 的视频信息。第一个报文从 00.00.00 开始,传播延时为 1s。理想状态下,第一个报文在 00.00.11 到达;第二个报文在 00.00.10 发出,在 00.00.21 到达;第一个报文在 00.00.20 发出,在 00.00.31 到达。但是,在图中,第一个报文从 00.00.00 开始,在 00.00.11 到达;第二个报文在 00.00.10 发出,在 00.00.28 到达;第三个报文在 00.00.20 发出,在 00.00.42 到达。那么第一个报文传播延时为 1s,第二个报文传播延时为 7s,第三个报文传播延时为 4s,这种不同的报文传播延时的不同叫作“延时抖动”。

对于实时的会话应用来说,从人们进行对话时自然应答的时间考虑,网络的单程传输延时应在 100～500ms 之间,一般为 250ms。在交互式多媒体应用中,系统对用户指令的响应时间也不应太长,一般要小于 1～2s。延时抖动将破坏多媒体的同步,从而影响音频和视频

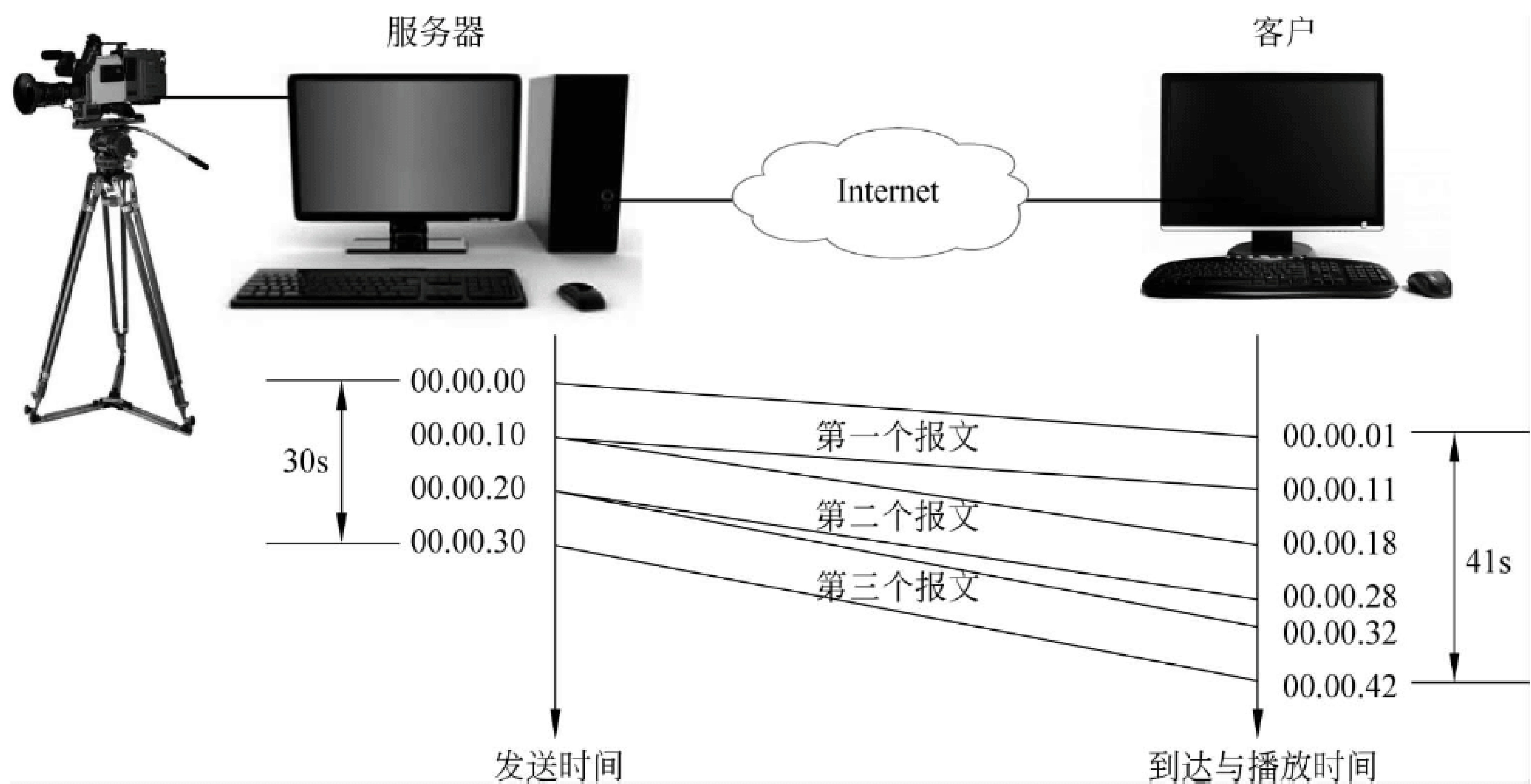


图 6-4　延时与延时抖动示意图

的播放质量。例如,数字的声音信号间隔的变化会使声音产生断续或变调的感觉。图像各帧显示时间的不同,也会使人感到图像停顿或跳动。人耳对声音的变化比较敏感。如果从一个人熟悉的音乐中删节很小的一段(例如 40ms),他立刻会感觉到。人眼对图像的变化就没有那么敏感。如果在熟悉的录像片中间删掉 1s(无伴音时)长的一段,我们未必能够感觉出来。因此,声音的实时传输对延时抖动的要求比较苛刻。考虑到网络性能与人的敏感度等实际情况,一般对不同应用给出以下的定量指标:对于经压缩的 CD 质量的声音,延时抖动一般应不超过 100ms;对于 IP 电话的语音信号,延时抖动不应超过 400ms;对于虚拟现实这类对传输延时有严格要求的应用,延迟抖动不超过 20～30ms。由于视频一般总是图像与伴音同步传送,需要从对伴音的要求去考虑对视频信号传输的延时要求:对于已压缩的 HDTV,网络延时抖动应不超过 50ms;对于已压缩的广播质量的电视,不超过 100ms;对于会议电视,网络延时抖动应不超过 400ms。表 6-3 给出了 IPTV 业务的端-端 QoS 要求的具体指标。我们将网络多媒体的这两个特点总结为:延时敏感与丢失容忍。

表 6-3　IPTV 业务的端-端 QoS 要求

业务类型	延时	延时抖动	丢包率	错误率
视频直播	1s	1s	0.01%	0.001%
视频点播	2s	2s	0.01%	0.001%
可视电话	150ms	50ms	0.01%	0.001%
视频会议	150ms	50ms	0.01%	0.001%
网络游戏	200ms	N/A	N/A	N/A

3. 网络多媒体数据传输方式

网络多媒体应用有两种基本方法:一种是下载后播放;另一种是采用流媒体方式,边下载边播放传输方式。最典型的下载后播放应用是 MP3。我们经常会从网上下载喜欢的歌曲,然后用 MP3 播放器播放。但是,对于网络技术最有挑战性的应该是流媒体的传输方式。目前,流媒体方式可以进一步分为三类:流式存储视频与音频、流式实况视频与音频、实时交互视频与音频。

1) 流式存储视频与音频

流式存储视频与音频最主要的特点是:边下载边播放。目前很多网站都能够提供流式存储视频与音频服务。

提供这种服务视频节目的网站将视频节目首先录好,并存储在服务器中。用户可以根据个人的兴趣向服务器提出服务请求。服务器响应的时间一般控制在1~10s。当服务器接受请求之后,用户端就开始接收服务器传送的视频数据,经过几秒或十几秒的启动延时即可进行观看。用户端一边播放,一边从服务器继续接收后续的数据段,直至完全接收位置。在播出期间,用户可以暂停、倒退、快进或者检索视频的内容。我们将这种技术称为流或流媒体技术。在这种方式中,视频数据在播放前并不下载整个文件,只将开始部分内容存入内存,后续的数据流随时传送随时播放,只是在开始时有一些延迟。尽管流式存储视频与音频可以满足一般用户连续播放视频节目的需求,但是它对端-端延时的要求上低于实况与交互式的多媒体应用。

2) 流式实况视频与音频

传统的电台广播室通过无线频道,例如天津音乐电台的频道是100.5MHz,采用的是调频FM方式,那么我们在开车时想听音乐,将车载收音机调到FM、100.5MHz时就可以收听到美妙的音乐。如果这个时间电台正在播"敖包相会",我们只能听到这个歌曲。同样,我们在家可以用电视机收看不同频道的节目,有时我们想看"2014年英超第38轮曼城对西汉姆的球赛"直播,可是已错过时间,现在正在播"2014年法网女单第一轮比赛"。那么,我们自然会想到:能否将电台、电视台的实况转播节目"搬到"Internet上,那么对于一些球迷或电视剧爱好者来说,他们就可以在去上班的地铁上、机场候机时,通过手机、iPad或笔记本收听到或收看到当天的球赛的实况转播或热播的电视剧。流式实况视频与音频应用可以实现我们收听、收看世界任何一个角落发出的实况无线广播或电视直播节目的愿望。目前,Internet上有数以千计的流式实况视频与音频电台与电视台。

由于是"实况"直播音频或视频节目,因此接收端的手机、iPad、笔记本一般不对接收到的音频或视频数据进行本地存储。为了有效地进行流式"实况"视频与音频播放与分发,服务器端一般要采取P2P或CDN方式实现多播,或者是采用独立的服务器到客户端的单播流传播方式。流式实况视频与音频方式的实时性要弱于实时交互视频与音频。用户从请求传输、播放到播放开始,大致要容忍最多几十秒的延时。

3) 实时交互视频与音频

典型的实时交互视频与音频应用是IP可视电话与网络视频。IP可视电话使人们在通话时能够看到对方图像,它不仅适用于家庭生活,而且还可以广泛应用于各项商务活动、远程教学、安防监控、医院护理、医疗诊断、科学考察等不同行业的多种领域。

IP电话又称为"VoIP(Voice over IP)"或"IP phone"。理解IP电话的概念,需要注意以下几个问题。

第一,我们一般将IP电话理解为:用脉冲编码PDM技术将人通话时的模拟语言信号变成数字语音信号,然后通过IP分组传输数字语音信号,在接收方将数字语音信号还原成模拟语音信号,实现双方语言交互的目的。这种情况相当于如图6-5(a)所示的情况。实际上,市场上已经出现了很多IP可视电话、网络视频与网络即时通信的应用软件,它们既能够实时传输会话双方的语音信号,还能够传输双方的图像与视频信号。

第二,IP 电话可以理解为固定电话、手机,以及各种类型的 IP 电话终端设备。通信方式可以是 IP 电话终端设备通过 Internet、IP 电话网关、电话交换网 PSDN 与固定电话的连接(如图 6-5(b)所示);可以是通过固定电话、电话交换网 PSDN、IP 电话网关、Internet、IP 电话网关、电话交换网 PSDN 与固定电话的连接(如图 6-5(c)所示)。同时,IP 电话可以理解为固定电话、手机通过 Internet、3G/4G 或 PSTN 互联的通信方式。

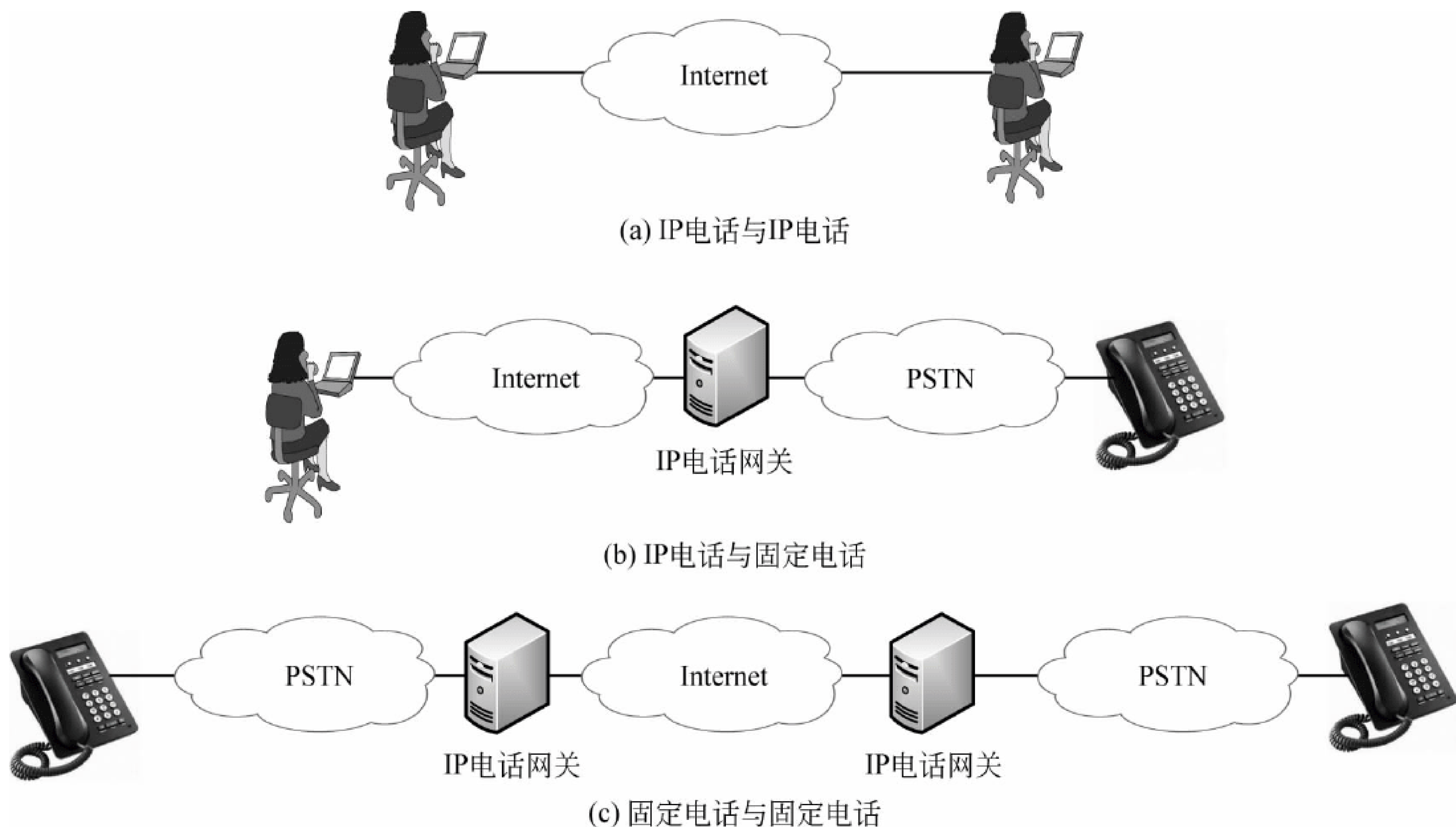


图 6-5　IP 电话的多种连接方式

第三,IP 电话网关实现 IP 网络与电话交换网之间的互联。IP 电话网关的作用主要有两个:一是在电话呼叫阶段与释放阶段完成电话信令的转换;二是在通话过程中实现语音编码标准的转换。

第四,传统的电话与 IP 电话在通信质量上差异比较大。因为传统的电话是通过线路连接的电话交换网传输,因此通话质量有保障。而 IP 电话则是通过只能提供“尽力而为”服务的 IP 分组交换网传输的,因此如何保证 IP 电话的通信质量是一个大问题。影响 IP 电话通信质量的因素主要有两个:通话双方的端-端延时与延时抖动,以及语音分组的丢包率。在实际应用中,人们发现:电话通信过程中端-端的延时不能大于 250ms,超过 250ms 通话双方就觉得不自然,而超过 400ms 就无法忍受。

随着多媒体网络应用的发展,为网络多媒体的一种通用、实时交互式应用的传输协议——实时传输协议(Real-Transport Protocol,RTP)与实时传输控制协议(Real-Transport Control Protocol,RTCP)应运而生。

问题 6-19:如何认识 RTP 的特点及其与相关协议的关系?

RTP 是由 IETF 的 AVT 工作组(Audio/Video Transport WG)提出的,最早的 RFC3550、RFC3551 文档定义了 RTP 与 RTCP,之前的 RFC1889 文档已经废止。RTP 已经成为 Internet 的正式标准,同时也成为 ITU-T 的 H.225.0 标准。

理解 RTP 的特点,需要注意以下几个问题。

1. RTP 的特点

1) RTP 运行在 UDP 之上

RTP 通常运行在用户空间,它位于 UDP 之上。如果从工作流程看,RTP 运行在用户空间,并且与应用层协议链接,因此它看上去更像是应用层的协议。而另一方面,它又是一个与具体应用无关的通用协议,它将应用层多媒体数据封装后,再利用 UDP、IP 及低层协议实现多媒体数据的传输。RTP 封装的多媒体信息可以是 PLC、GSM 与 MP3 的数字语音数据流,也可以是 MPEG 与 H.263 视频流。

RTP 是一个框架,它包含传输实时应用数据流的共同特性。RTP 协议只包含实时多媒体应用的一些共性的功能,它并不对多媒体数据流做任何特殊的处理,只是通过与 RTP 协同工作的 RTCP,向应用层提供在当前网络条件之下,能够尽可能高地提高服务质量的相关信息。

当应用程序开始一个 RTP 会话时,将使用两个端口,一个给 RTP,另一个给 RTCP。同时需要注意的是,不像其他的应用层协议,都是分配一个熟知端口号,而 RTP 会话需要在临时端口号的 1025～65 535 之间,选择一个未使用的偶数 UDP 端口号。例如,RTP 选择的端口号为 1210,那么属于同一会话的 RTCP 使用加 1 的奇数端口号,即 1211。因此从 IP 与协议体系的角度看,它应该是在应用层之下,UDP 之上插入,专门用于对实时性有要求的网络应用的传输层协议。

图 6-6 给出了引入 RTP/RTCP 之后的网络层次结构模型。

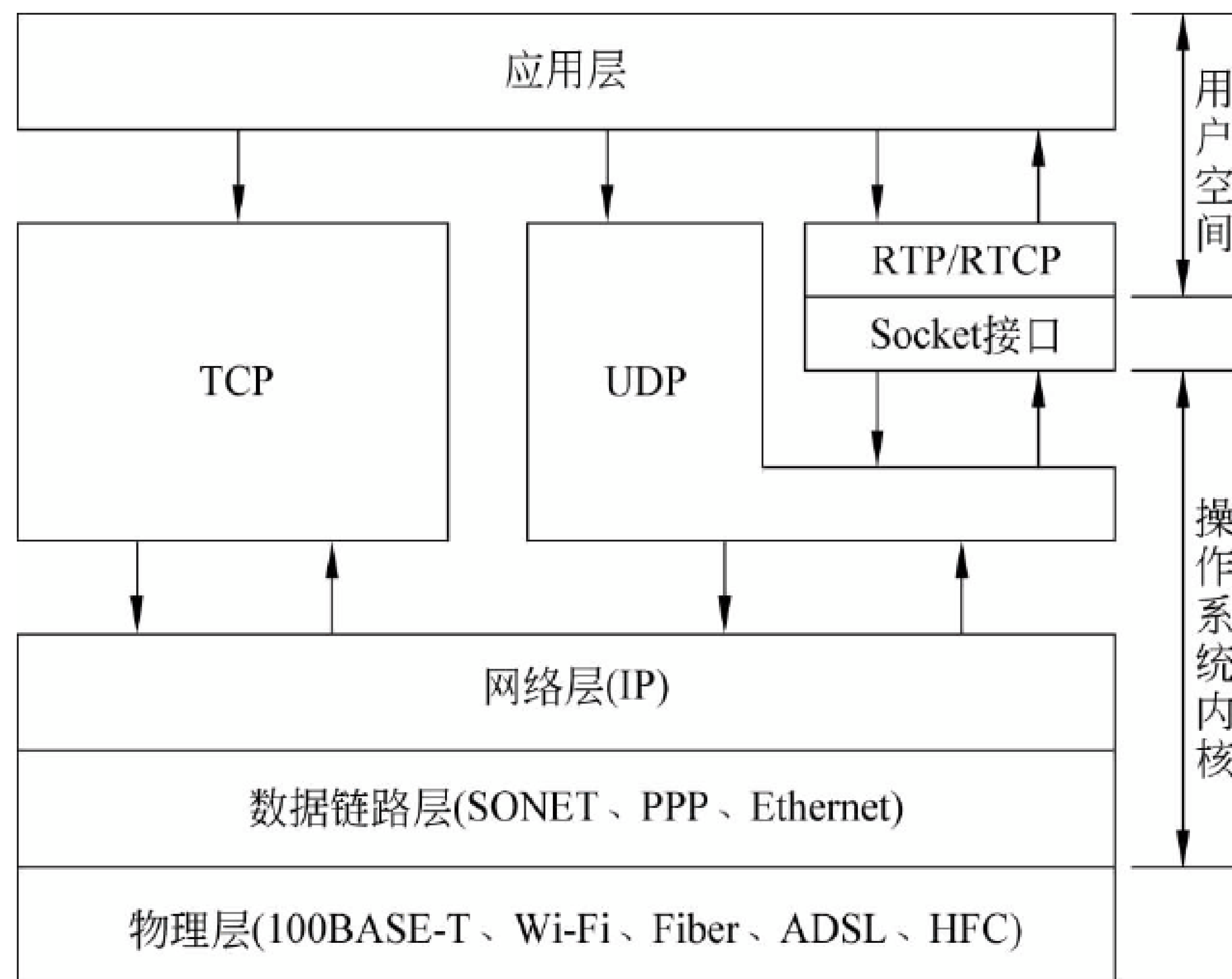


图 6-6 RTP 层次结构模型

2) RTP 提供端-端的传输服务

多媒体数据是由音频、视频、文本与其他可能的数据流组成。这些数据流送到 RTP 库。RTP 库软件将按照音频、视频、文本数据流之间的关系压缩编码后复用到 RTP 报文(RFC3550 使用的是"RTP packet"),加上套接字(Socket),使用操作系统内核中 UDP 软件,封装成一个新的 UDP 报文。目的主机将接收到的 RTP 数据包封装的多媒体数据传送到应用层。应用层的播放器负责播放多媒体节目。图 6-7 给出了 RTP 与 UDP、IP、Ethernet 协议数据单元之间的关系示意图。

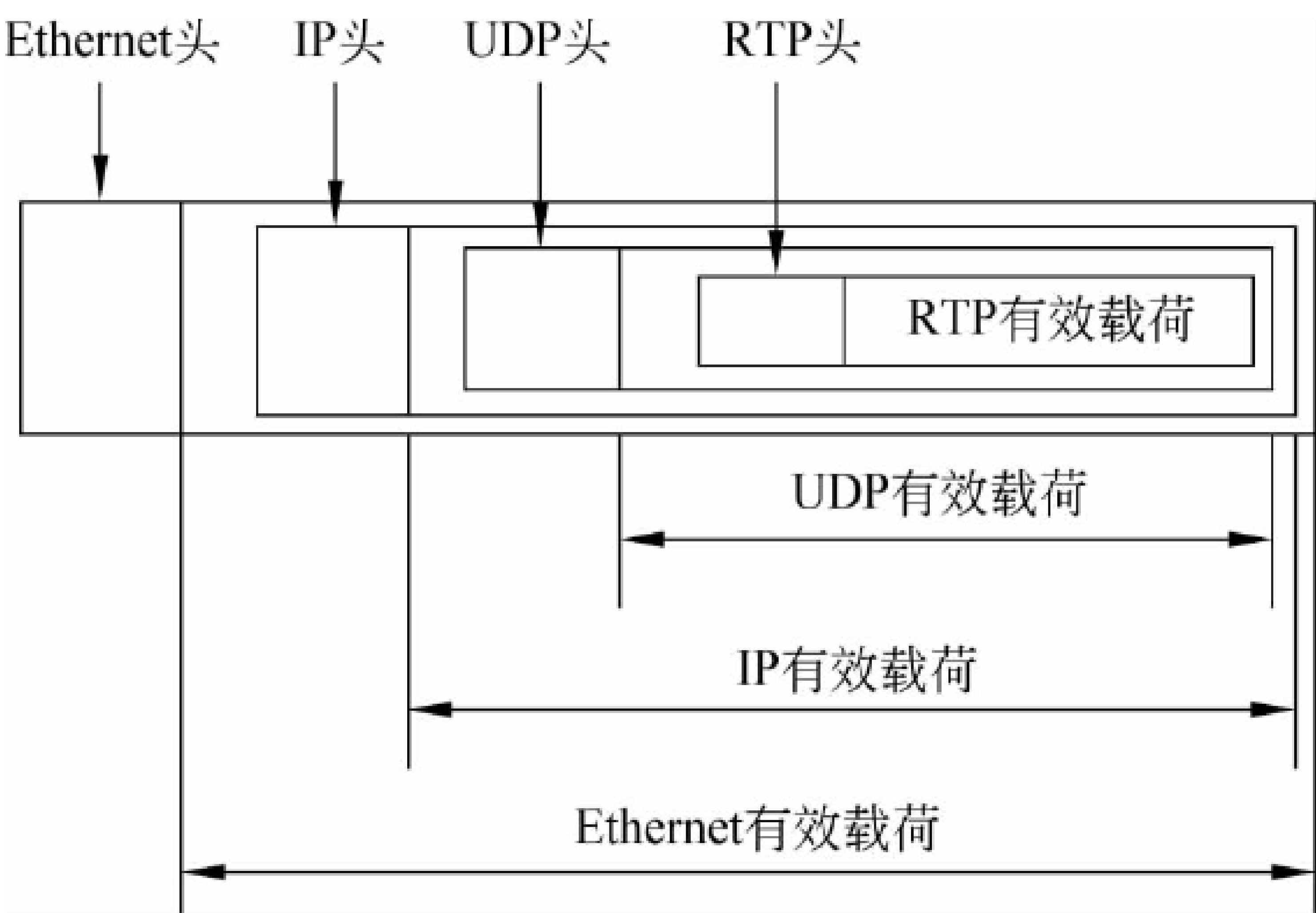


图 6-7　RTP 与 UDP、IP、Ethernet 协议的关系

UDP 数据报文段封装在普通的 IP 分组中传输，传输路径中的所有路由器不会对该分组提供任何特殊的服务。RTP 不强调资源预留协议（RSVP）的支持，RTP 为应用层实时应用提供端-端的传输服务，不提供任何 QoS 保证。

2. RTP 的结构

RTP 报头结构如图 6-8 所示。

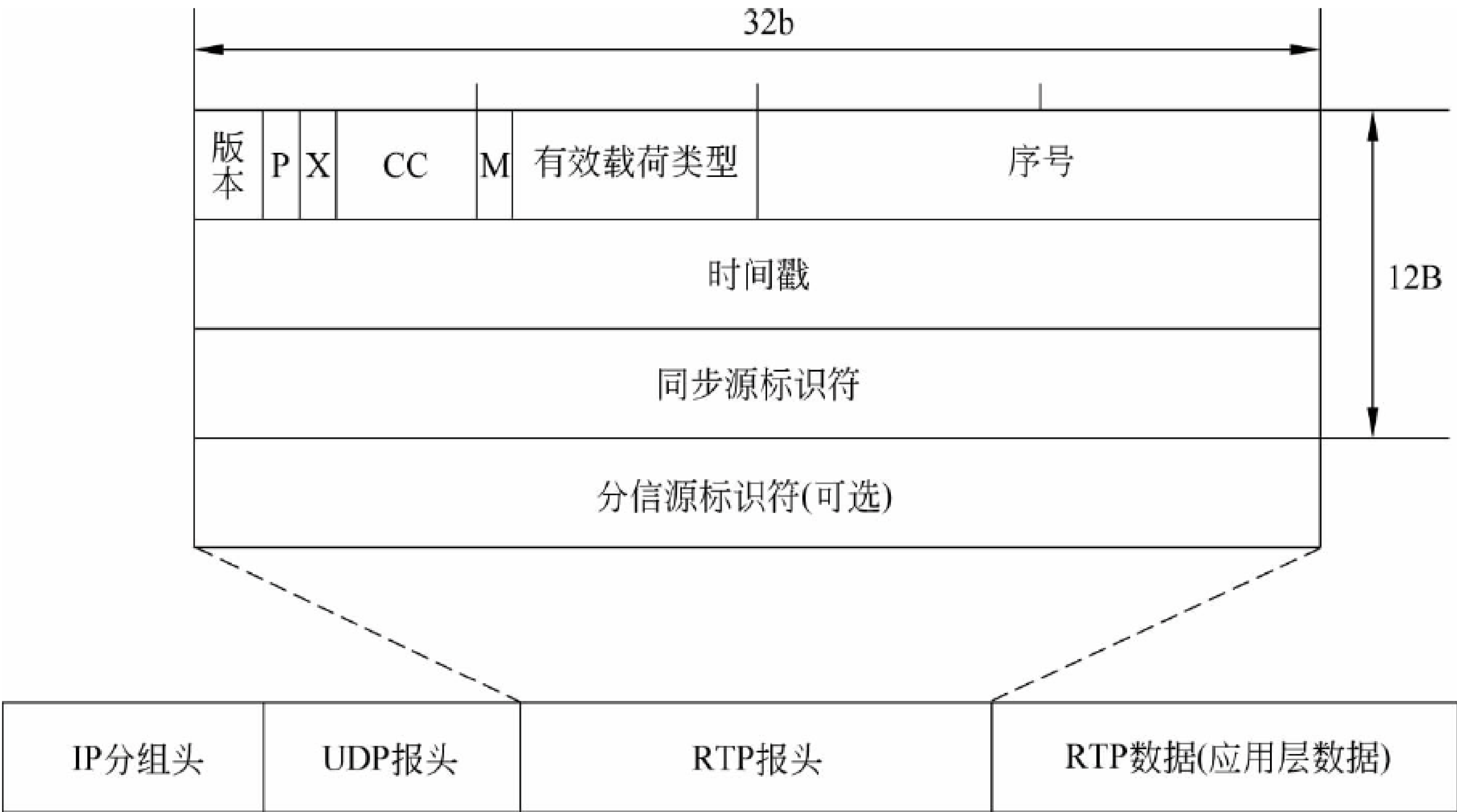


图 6-8　RTP 报头结构

RTP 报头由 12B 固定长度的 RTP 报头与可选的分信源标识符组成。长度为 12B 的固定 RTP 报头包括以下字段。

1）版本

版本字段长度为 2b，目前使用的是版本 2。

2）填充

填充（P）字段长度为 1b。在某些特殊的情况下需要对应用层数据进行加密，这就要求每一个数据块有确定的长度，必须是 4B 长度的整数倍。在有填充字节的情况下，填充位 P=1。在数据部分的最后一个字节值用来表示所填充的字节数。

3) 扩展

扩展(X)字段长度为 1b。X=1 表示 RTP 报头之后有扩展报头。RTP 很少用扩展报头。

4) 参与源

参与源(CC)字段长度为 4b,表示一次会话最多有 15 个参与源。

5) 标记

标记(M)字段长度为 1b。M=1 表示该 RTP 报文有特殊意义。例如,应用程序可以用这一位表示传输视频流的每一帧的开始,也可以用于指示视频流传输的结束。

6) 有效载荷类型

有效载荷类型字段长度为 8b。目前已经定义的音频与视频类型(编码格式)如表 6-4 所示。

表 6-4　有效载荷类型

有效载荷类型编号	音、视频类型	有效载荷类型编号	音、视频类型
0	PCMμ 律音频	15	G. 728 音频
1	1016 音频	26	运动 JPEG 视频
3	GSM 音频	31	H. 261 视频
7	LPC 音频	32	MPEG1 视频
9	G. 722 音频	33	MPEG2 视频
14	MPEG 音频		

7) 序号

序号字段长度为 16b,用来给 RTP 报文编号。一次会话的第一个 RTP 报文编号是随机获得的;每个后续的报文序号加 1。接收方根据序号来判断 RTP 报文是否丢失或乱序。

8) 时间戳

时间戳字段长度为 32b,用来指出 RTP 报文之间的时间关系。在一次会话开始时,第一个 RTP 报文的时间戳初始值是随机产生的。RTP 并没有规定时间戳的粒度,它取决于有效载荷的类型。例如,对于采样时钟为 8kHz 的语音信号,每隔 20ms 产生一个数据块,每个数据块包含 160 个样本(8000×0.02=160)。发送方每发送一个 RTP 报文,其时间戳的值就增加 160。接收端可以根据时间戳值,确定如何还原数据块的时间,以消除延时抖动。同时,时间戳也可以起到视频应用中的音频与图像同步的作用。

9) 同步源标识符

同步源标识符(Synchronous SouRCe identifier,SSRC)字段长度为 32b,用来表示 RTP 流的来源。如果一次会话只有一个源端,那么 SSRC 字段的值就表示这个源端。如果有多个源端,那么混合器就是同步源端,而其他源端是参与源。通过 SSRC 字段可以将多个数据流复用在一个 UDP 报文中;同时也可以从一个 UDP 报文将多个数据流分离出来。

10) 参与源标识符

参与源标识符(Contributing SouRCe identifier,CCRC)字段长度为 32b,用来标识参与源的源端。从长度为 4b 的参与源(CC)字段可以知道,一次会话参与源的数量最多为 15 个。

问题 6-20：如何认识 RTCP 与 RTP 的关系？

理解 RTCP 与 RTP 的关系，需要注意以下几点。

第一，RFC3550 文档实际上定义了两部分的内容。一部分是用于传输多媒体数据流的 RTP，另一部分定义的是实时传输控制协议 RTCP。

第二，RTCP 与 RTP 是相互配合的关系。RTP 与 RTCP 可以同时在一个多媒体应用中使用，都封装在 UDP 报文中传输。

第三，RTP 报文的有效载荷封装音频、视频流，而 RTCP 报文不封装任何音频、视频数据流。

源端可以利用 RTCP 报文同步一次会话中不同的媒体流。例如，在一次视频会议应用中，每个源端都产生了两个独立的媒体流，一个用于视频传输，一个用于音频传输。需要将这些 RTP 报文头中的时间戳与视频、音频采样时钟建立关联。由于源端发出的 RTCP 报文包含与之相关联的 RTP 报文流的时间戳与真实时间，接收端就可以通过 RTCP 报文提供的关联，同步视频与音频的播放。

表 6-5 给出了 5 种 RTCP 报文的类型与功能。RTCP 报头中有一个长度为 8b 的报文类型字段，不同的报文类型字段值表示不同类型的 RTCP 报文。例如，报文类型字段值为 200 表示发送端报告的 RTCP 报文。

表 6-5　RTCP 报文的类型与功能

报文类型字段	英文缩写	功　　能
200	SR	发送端报告
201	RR	接收端报告
202	SDES	源点描述报告
203	BYE	结束
204	APP	特定应用

1. 源端报告(SR)

发送端与接收端的一次会话包含着很多的 RTP 流。发送端每发送一个 RTP 流时，就发送一个源端报告(SR)报文。SR 报文包括：

(1) 该 RTP 流的同步源标识符 SSRC。

(2) 该 RTP 流最新产生的 RTP 报文的时间戳与绝对时间(又称为“墙上时钟时间”)。

(3) 该 RTP 流包含的报文数。

(4) 该 RTP 流包含的字节数。

绝对时间对于多媒体传输是非常重要的。因为传输一个视频信号时实际上需要同时传输音频流与图像流，播放视频节目时可以用 RTP 报文的时间戳与绝对时间实现图像与语音流的同步。

2. 接收端报告

接收端每接收一个 RTP 流时，就会产生一个接收端报告(RR)报文。RR 报文包括：

(1) 接收到的 RTP 流的同步源标识符 SSRC。

(2) RTP 流的报文丢失率。

(3) RTP 流最后一个 RTP 报文的序号。

(4) RTP报文到达时间的延时抖动。

接收端可以使用RTCP报文,周期性地向源主机反馈与服务质量相关的统计数据。源主端可以根据RTCP报文反馈的信息,了解网络当前的延时与延时抖动、丢包率,来决定数据传输速率。如果网络通信状态好,则源端可以动态地改变编码算法,以提高多媒体信息的播放质量。例如,如果网络延时、延时抖动与丢包率都很低,那么源端可以将语音编码从MP3切换到需要占用更多带宽的8位PCM,或者是切换到增量编码方式,从而在当前条件下提供尽可能好的服务质量。

3. 源点描述报告(SDES)

源端周期性以多播方式,通过发送源点描述报告(SDES)报文,给出会话参与者的规范名(Canonical Name)。规范名是会话参与者电子邮件地址字符串。

4. 结束

结束(BYE)报文用来关闭一个数据流。在视频会议应用中,一个源端可以用结束报文宣布退出这次会议。

5. 特定应用

特定应用(APP)报文用于应用程序定义一种新的RTP报文类型。

对于一个规模较大的组播应用,RTCP报文占用网络带宽可能变得很大。为了防止这种现象出现,所有的发送RTCP报文的节点自适应地调节自己发送RTCP报文的速率,使得起到控制作用的RTCP报文不要过多地占用网络带宽而影响RTP报文的传输。通常控制RTCP会话的报文占用网络带宽不超过5%。假如有一个发送方正在以2Mbps的速率发送视频流,那么该节点的RTCP报文占用网络带宽必须低于100kbps。具体实施时,将这个带宽的75%(75kbps)分配给接收方,剩余的25%(25kbps)留给接收方。如果在多播情况下有n个接收方,那么每个接收方能够用于发送RTCP报文占用网络带宽只能控制在75/n(kbps)之内。然后,每个接收方将根据能够使用的带宽,以及发送的RTCP报文平均长度,计算出发送RTCP报文的周期。

问题6-21:如何认识容迟网DTN技术的研究背景?

了解容迟网技术,对于理解物联网应用很有帮助。教师用书增加了一部分有关容迟网技术的内容。理解DTN研究的背景,需要注意以下几个问题。

(1) 在设计Internet应用层协议工作原理时都是有一个假设:在一次进程通信过程中,源端与目的端之间一定要保证“持续”的TCP传输连接。如果不能保证TCP“持续”连接,那么分布式进程通信失败,网络服务不能实现。我们大量使用的Internet服务,如Web、E-mail、FTP都是建立在这个“假设”的基础之上。目前存在着很多应用,如低地球轨道(Low-Earth Orbit,LEO)卫星通信网、星际网络、水下无线传感器网络、地下无线传感器网络、军用无线传感器网络、GPS网络与无线车载网VANET等网络应用,实际上是运行在一个复杂的“受限网络”之上,这个“假设”都是无法保证的。

(2) 1998年,美国NASA开始了星际互联网(InterPlanetary Internet,IPN),也称“深空网络”的研究。星际互联网研究的基本目标是:让地球和距离很远的太空船之间的数据通信,能够简化到像地球上Internet中的两个节点通信一样的方便。

从表面上看,这个研究有一点儿“太困难”,但是恰恰是这个灵感,推动了容迟网(Delay-Tolerant Network,DTN)技术的研究与发展。“容迟网”又称为“中断容忍网络(Disruption-

Tolerant Network，DTN）”。

NASA 的研究人员后来成为 Internet 的 IPNSIG 工作组。但是 IPNSIG 遇到的一个问题是：目前还没有这样一个星际网络可以进行实验，于是有一部分人开始研究如何将 IPN 的概念运用到陆地网络中。为此，IETF 成立了新的 DTNRG 工作组，研究更加通用的容迟网络体系结构、技术与标准。

（3）DTN 的主要特点。

① 长延时。

例如在星际通信中，地球与火星距离最近时，光传播需要 4min 时间，而距离最远时的光传播时间会超过 20min。在 Internet 中，传播时间一般以毫秒计算，因此如此长的延时，传统的 TCP/IP 是无法适应的。

② 间歇性连接。

当空间节点之间受到其他星球的阻挡，当地面移动节点与基站之间有建筑物阻挡，当地空卫星移动出卫星地面站接收范围，当无线车载网的节点之间受到其他车辆阻挡时，都会造成节点之间端-端连接的间歇性断开。在一辆经常往返的公共汽车上安装射频通信系统就可以被用做信息存储和转发的工具。当这辆公共汽车从一个地方开到另一个地方时，它可以在附近的客户机和它将要去的地方的远程客户机之间提供信息交换服务。这些端-端连接的中断可以有一定规律，也可以是随机的。但是，传统的 TCP 是不支持的。

③ 低信噪比和高误码率。

无线、移动与长距离传输会导致接收信号的低信噪比与高误码率。在 Internet 中光纤传输的误码率可以达到 10^{-12}～10^{-15}，而太空通信中的误码率甚至可以达到 10^{-1}。这种低信噪比、高误码率会极大地影响接收端对信号的解码和恢复，造成 TCP 连接的非正常中断，使得网络系统不能正常工作。

④ 不对称数据速率。

在特殊的网络应用中，数据传输的双向速率经常是不对称的。在完成空间任务时，双向数据速率比可以达到 1000∶1 甚至更高。这也是传统的 TCP 设计时没有考虑到的情况。

⑤ 节点资源的限制。

应用于太空、水下、战场、救灾现场与环境监测等环境中的无线传感器节点受体积和重量的限制，电源与计算、存储资源非常有限，它不可能像办公环境中有电源供应保障的 PC 一样，有足够的电源、计算与存储资源与网络带宽。无线传感器节点经常会因电池能量的耗尽而停止工作，其他节点要重新计算路由。无线传感器节点经常会因为节省电能而处于休眠状态，这时只有其他的节点唤醒它或休眠时间结束，它才能进入加入到无线自组网中的状态。在这种情况下，端-端连接也会经常中断。

（4）从 2002 年开始，针对间歇性通信与长延时的网络消息交换（Message Switching）技术需求的“延迟容忍网络”引起了学术界的重视。人们一直在开展 DTN 协议、模型与应用的研究。2003 年，Kevin Fall 观察到这些星际 Internet 的思想可以应用于地球上一些具有间歇连接特征的网络应用，采用类似于电子邮件的节点存储、延时转发，最后数据被中继到目的节点，而不是采用传统的由路由器进行分组交换来完成，在此基础上 Kevin Fall 提出了 DNT 体系结构模型。

（5）当 IDC、云计算、大数据应用趋势日趋明晰之后，人们也发现了 DTN 在 IDC 大规模

数据存储与备份中的应用前景。由于一个大型的数据中心IDC可能在世界各地设立有多个分中心,多个分中心之间需要许多太字节量级的复制、存储。运营商希望在非高峰时间传送这些大块的数据,以便均衡链路的负荷,减少费用,同时这种非实时的数据备份可以容忍有一定的时间延迟。例如,当夜深人静的时候,网络的利用率低,这个时间用来在IDC之间传输大量的数据。尽管这些IDC分布的地区较大,有的地方是夜间,而有的地方已经天亮了。例如,波士顿与帕斯的IDC的非高峰期网络带宽在时间上很少有重叠,但是由于DTN模型允许在数据传输过程中进行存储与延时,因此设计者可以考虑将数据实现集中存储到只有6小时时差的阿姆斯特丹,直到阿姆斯特丹与帕斯之间的非高峰网络带宽可用时,再将数据集中发送到帕斯。Laoutaris等研究的结果表明:这样做可以用较小的费用获取很大的带宽容量;与传统的方法相比,DTN模型通常将容量扩大一倍。目前,研究DTN技术的三个主要机构如下。

① 星际互联网(Interplanetary Networking,IPN)研究组。

② IRTF建立的DTNRG(DTN Research Group)研究组。

③ 美国国防部高级研究计划局(DARPA)。

问题6-22:什么是DTN体系结构?

DTN研究的基本思路是:基于消息交换的体系结构,容忍低可靠性、大延时的链路。2007年发表的RFC4838对DTN体系结构进行了说明。基于消息交换的DTN体系结构如图6-9所示。

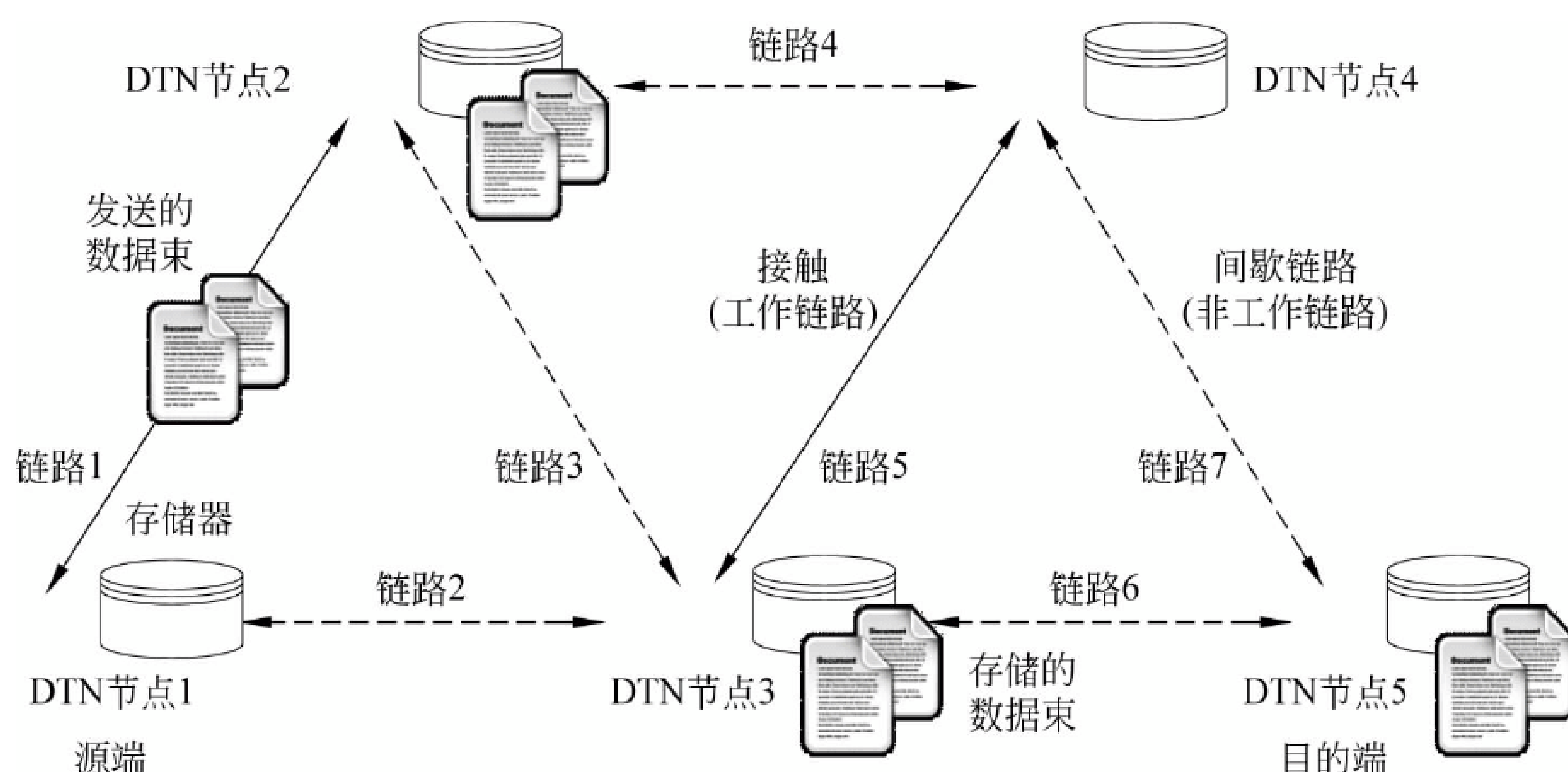


图6-9 DTN体系结构示意图

理解DTN体系结构需要注意以下几个概念与特点。

1. 数据束

按照DTN的术语,一条“消息”称为一个“数据束”。DTN节点都配置了磁盘或闪存等存储介质。它们将数据束存储起来,直到链路变得可用时,再转发数据束。图中有5个DTN节点,其中DTN节点1为源端,DTN节点5为目的端。

2. 链路与接触

由于DTN中链路的工作呈现间歇性的特点,因此链路根据它的工作状态分为“工作链路”与“非工作链路”两种。图中有两条工作链路(链路1与链路5),5条间歇链路(链路2、

3、4、6 与 7）。一条工作链路称为一次“接触”。图中显示，通过工作链路，数据束已存储到 DTN 节点 2 与节点 3。节点 2 与节点 3 将利用下一次“接触”的机会，通过工作链路将数据束转发下去，直到目的端的节点 5。

3. DTN 节点与路由器之间的区别

表面上看 DTN 节点接收、存储、转发数据束，路由器也是接收、存储、转发分组，但是两者有很大的区别。第一，传统 Internet 中的路由器转发 IP 分组的延时一般都控制在毫秒或秒的量级，而 DTN 节点接收数据束之后需等待较长时间，甚至是几个小时的时间才能够转发数据束。例如，IDC 中心之间利用网络非高峰时间传送大块的数据需要根据网络运行状态来决定等待时间；利用飞机转发数据束，需等到飞机到达之后才能进行；利用公共汽车转发数据束，需等待公共汽车到站之后才能进行。第二，传统的路由器一般是不移动的，但是很多 DTN 应用中，作为 DTN 节点的卫星、飞机、汽车、手机都是移动的。而且我们恰恰是利用的 DTN 节点的移动性来以最小代价实现数据束的“存储、携带和转发”。

问题 6-23：如何认识 DTN 协议体系模型与数据束协议的特点？

认识 DTN 协议体系模型与数据束协议的特点，需要注意以下几个问题。

1. DTN 协议体系模型

DTN 网络协议体系结构分为高层、DTN 层与低层，其体系结构模型如图 6-10 所示。

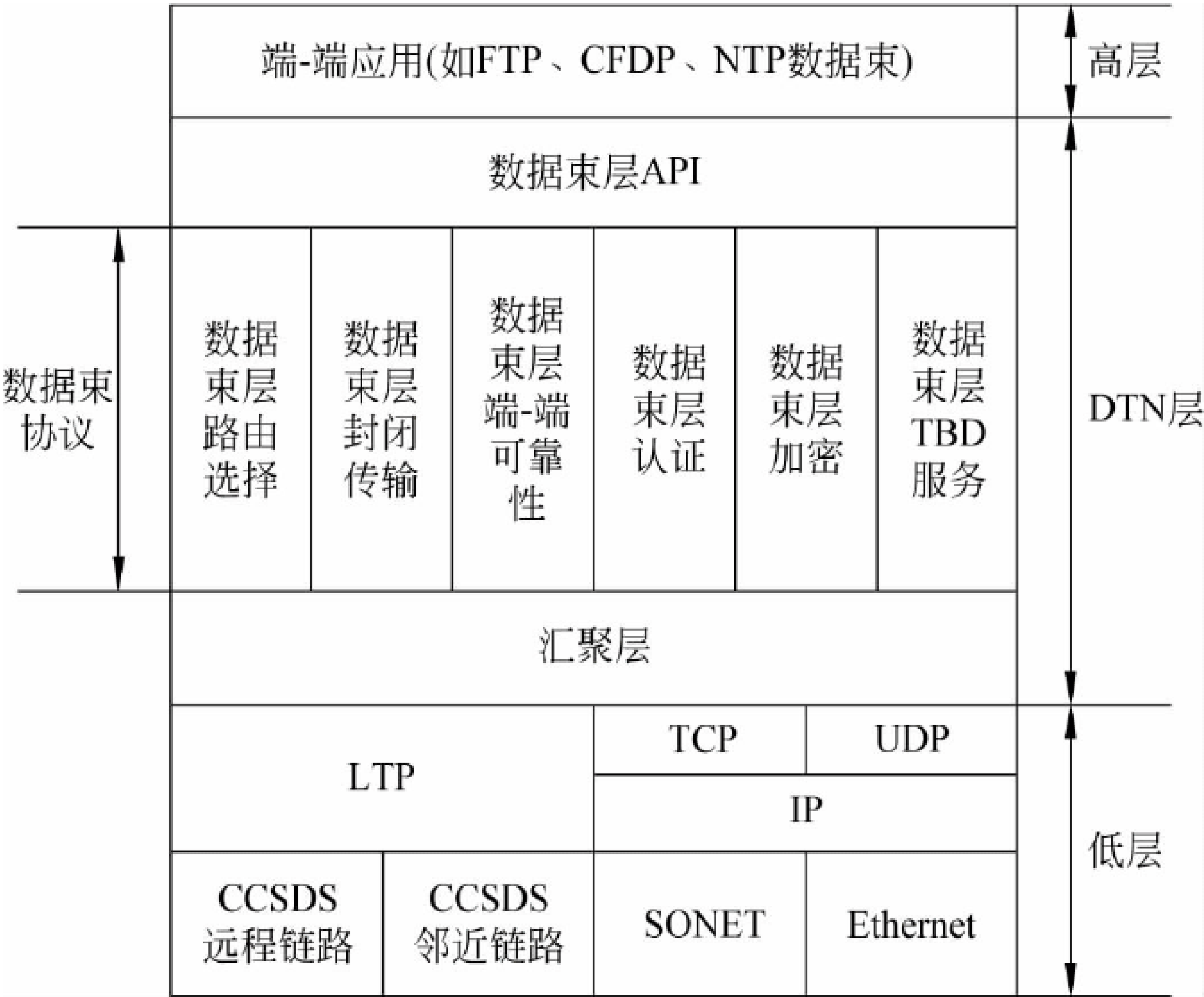


图 6-10　DTN 体系结构示意图

1）应用层

DTN 网络协议体系结构的高层是指应用层，DTN 应用层由适应各种不同网络应用系统的应用层协议组成。

2）DTN 层

DTN 层包括数据束协议层与汇聚层。数据束协议负责接收应用程序的消息，通过“存储-携带-转发”的方式，将应用程序的数据作为一个或多个数据束转发到目的 DTN 节点。

3) 低层

DTN网络协议体系结构的低层是指TCP、UDP或其他的传输层协议,以及相应的数据链路层与物理层协议。

2. DTN层的特点

理解DTN层的特点,需要注意以下几个问题。

第一,为了向应用层屏蔽系统由于DTN网络要使用低可靠性、高延迟的链路,采取"存储-携带-转发"数据的过程,设计者在应用层与传统的互联网传输层之间设计了数据束协议(Bundle Protocol,BP)。该协议接收来自应用层的应用进程的消息,并通过"存储-携带-转发"操作,将这些数据作为一个或多个数据束发送到目的DTN节点。数据束协议是运行在TCP/IP之上,数据束协议可用在DTN节点之间的每个间歇性通信中,将数据束从一个DTN节点转发到另一个DTN节点。数据束协议属于传输层协议,但它为很多不同的应用提供传输服务,因此它需要覆盖整个DTN网络。

第二,数据束协议可以运行在TCP或UDP之上,也可以运行在其他的传输层协议之上。由于数据束协议是不变的,而采用的传输层协议可以是TCP、UDP或其他的传输层协议,因此DTN体系结构在数据束协议层与传统的传输层之间增加了汇聚层。汇聚层的主要功能是实现TCP/UDP或其他传输层协议报文,与数据束协议报文格式之间的转换。它向数据束协议屏蔽了TCP、UDP或传输层协议格式、功能上的差异性。如果数据束协议运行在TCP/IP之上,那么TCP/IP就可以在DTN节点之间的每一次"接触"中,将数据束从一个DTN节点转移到另一个DTN节点。

第三,尽管DTN应用程序可以应用TCP,但是它也可以使用UDP或其他的传输层协议。很显然,在星际网络中链路传播延时是很大的,例如,地球与火星之间的数据传播往返的延时就可以达到20min,因此在这种情况下使用TCP确认、重传协议是无法工作的。在一些资源受限的无线传感器网络中,也必须使用简化的TCP。在这些应用中,必须考虑使用其他的传输层协议,或者是根据具体的进程通信特点,去研究新的传输层协议。针对不同的传输层协议需要设计不同的汇聚层。DTN汇聚层使得数据束协议具有适应不同应用的可扩展性。当然,DTN网络体系结构中的汇聚层协议与物联网体系结构中的汇聚层功能不同,但它为我们研究复杂网络体系结构提供了一种有益的启示。

3. DTN数据束协议

理解DTN数据束协议,需要注意一个问题:DTN是一种新兴的网络,实验网络也只是针对一些特殊领域的应用,目前IETF还没有推出标准的DTN数据束协议,DTN数据束协议仍处于专用网络协议阶段。图6-11给出了2007年11月发布的RFC5050"Bundle Protocol Specification"文档中定义的数据束协议的消息格式。

每个数据束消息是由主块、有效载荷与可选块组成。主块是消息的头部,有效载荷封装了数据,可选块可以携带有关安全的一些参数。

主块是由"版本""标志""目的端""源端""托管"等9个字段组成。下面选择其中几个重要的字节加以说明。

1) 主块的"标志"字段

"标志"字段包括状态报告(7b)、服务类型(7b)、其他(6b)。"状态报告"可以理解为"块处理控制标志",表示该数据束能不能被处理、复制、删除,以及是不是最后一个字段。源端



图 6-11　数据束协议的消息格式

可以用“服务类型”来表示数据束的优先级。“其他”字段值用来表示：有效载荷是不是数据流中的一个片段，有效载荷是不是管理程序的记录，数据束能不能分片，以及有效载荷是不是对应用程序请求的确认等。

2）主块“目的端”“源端”与“托管”字段

在主块中除了两个地址字段“目的端”“源端”之外，还有一个“托管”字段。

理解这几个字段的意义与作用，需要注意以下几个问题。

第一，在 Internet 中，源节点可以理解为“托管”节点，因为如果数据字段没有传送到目的节点时，源节点要负责重传。DTN 使用“托管传输”机制来处理这个问题。在数据束转发过程中，另一个接近接收节点的节点可以承担起“托管”节点的功能。

第二，DTN 标识符不是 IP 地址，它使用自己定义的、类似于 URI 的高层的标识符，以实现电子邮件与软件更新那样的应用程序级的路由寻址。

第三，DTN 标识符的编码方法是一个需要研究的问题。因为这对于节点计算、仓储与电能受限的无线传感器网络与某些物联网移动终端，用尽可能短的，但具有唯一性的节点标识符是至关重要的。DTN 标识符可以根据具体的应用场景，选择使用 MAC 层地址、网络层地址、网络标识符、资源标识符。也有一种研究是提出可变长度字典，它类似于 RFID 中对象名字服务 ONS 的方法，在数据束标识中使用简单的数字编码，通过字典可以“还原”出数据束的详细信息。

3）主块的“创建”“生存期”字段

“创建”字段用于表示携带的数据束的创建时间。“生存期”字段值表示数据束在网络中传输的最长时间。当数据束在网络中转发的时间已经超过“生存期”的数值，该数据束将被删除。

4)"块长度"字段

由于主块与有效载荷都有可变长度字段,因此主块与有效载荷都需要设置"块长度"项。

实践证明,RFC5050定义的数据束协议(BP)在推动容迟网络研究与应用发展中起到了重要的作用,但是在可靠性、错误检测、时间同步机制、网络安全、网络管理等方面仍需要改进。

问题6-24:DTN技术在星际网络中有哪些应用?

空间网络中开始应用了DTN技术。了解DTN技术的应用与星际网络的发展,对于进一步理解DTN协议的研究,以及计算机网络技术的最新发展很有帮助。

1. 目前DTN在星际网络中的应用研究的现状

图6-12给出了DTN技术用于空间网络的示意图。数据束是低地球轨道卫星LEO作为灾害监控星座卫星的一部分,记录的是地球图像。地球采集点收集LEO卫星传送的图像数据。如图6-12所示,卫星在绕地球旋转时只能间歇性地接触到三个地面站。三个地面站在每一次接触时接收卫星发送的数据束,然后地面站通过陆地网络,将接收到的数据束发送到采集点。采集点再将从各个地面站传送的数据束汇聚成完整的地球图像数据。在整个过程中,地面站在一次接触中,集中所有的带宽用于接收下行的图像数据。在完成一次"接触"之后,再在适当的时间利用陆地网络,转发接收到的图像数据。

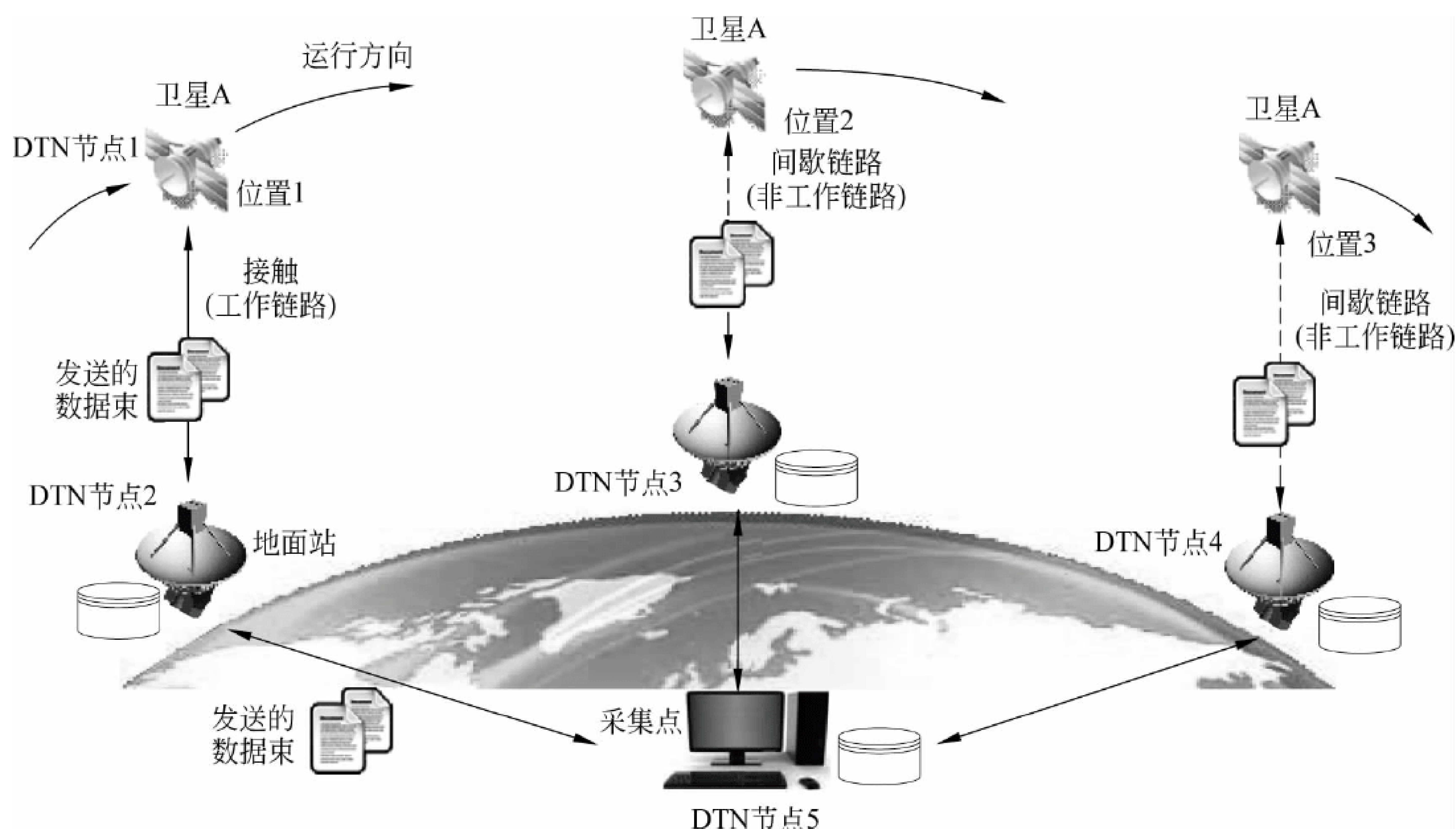


图6-12 DTN在LEO卫星通信网中的应用

美国NASA/JPL于1998年起开展"星际互联网(IPN)"的项目研究。图6-13给出了未来火星任务深空测控通信环境(Deep Space Communication Environment for Future Mars Exploration)示意图。它包括地面链路、地球轨道链路、星际主干链路、行星(如火星)轨道链路与行星表面链路。这些链路的特性都不相同。NASA空间互联网的基础设施是由主干网络、接入网络、航天器之间网络、近距无线网络等结构单元组成。地球上的Internet通过一个超长距离的无线链路,与深空骨干网相互连接在一起。

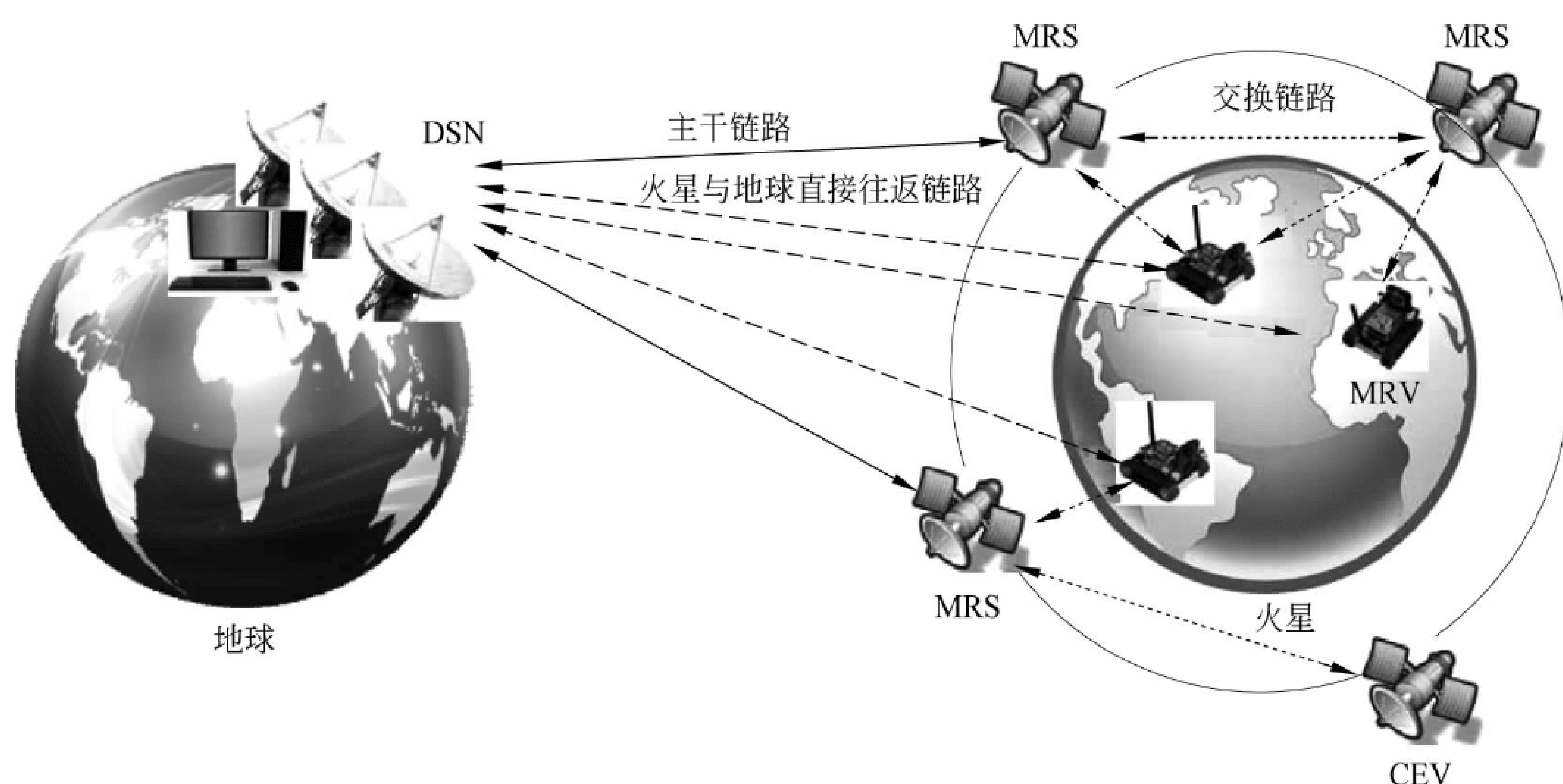


图 6-13　未来火星任务深空测控通信环境示意图

2004 年,NASA 成立了空间通信体系结构工作组(SCAWG),对直到 2030 年都能适应的 NASA 空间探索和科研任务的空间通信体系的概念、结构以及远景战略开展研究。

2008 年 11 月,美国宇航局喷气推进实验室(JPL)利用 DTN 技术在地球和 3200 万千米以外的航天器之间传送了一批空间图像。整个实验网络共设置了 10 个节点,其中一个是太空探测器,其他 9 个位于地面。他们用这样的网络环境模拟当太空探测器飞行到某颗行星背后,或在发生太阳风暴时,通信网络就会发生故障,研究 DTN 数据传输过程中的各个关键环节可能存在的问题。

2. 太空路由器的研究

作为太空 Internet 路由(IRIS)研究计划的一个重要内容,由 Cisco 公司制造的基于 IP 协议的太空路由器,随 Intelsat 的 IS-14 商用卫星于 2009 年 11 月 23 日一起发射升空。在传统的卫星通信模式中,两颗卫星之间传输数据必须要通过中继卫星或通过地面站。在卫星上安装太空路由器,则可在相邻卫星之间直接进行通信,省去了回传地面的往返时间,降低了网络传播延迟,可以大大提高卫星通信效率、降低成本。因此美国军方积极推动这项研究计划,并从 2010 年 1 月开始,美国国防部进行了为期三个月的演示,以评价太空互联网研究的军事应用价值。

实际上 IS-14 商用卫星上安装了两台 IRIS 路由器,其中一台作为备份。同是作为备份,每一路由器上还安装了两个独立的调制解调设备。每一台路由器与三组转发器组相连,每组由六十多个转发器组成,路由器的吞吐量可以达到 100Mbps。路由器采用了 Cisco 操作系统(IOS)软件,并利用 IPSec 协议,以提高数据传输的安全性。太空路由器的每一个组件都可经受大量射线的考验,预计使用寿命为 15 年。整套路由器设备的尺寸大约为 24 英寸×18 英寸×18 英寸。太空路由器外形与安装到卫星的工作照片如图 6-14 所示。

2010 年 12 月 22 日,俄罗斯宣布正计划构建太空网络,目的是支持航天器之间的联络,保障俄罗斯偏远地区的通信,实现在地球上任何地点都能对航天器进行控制的目标。

3. 深空通信网络协议体系研究的基本思路

“深空通信 DTN 应用研究”一文中对 DTN 技术在深空网络中的应用进行了较为全面

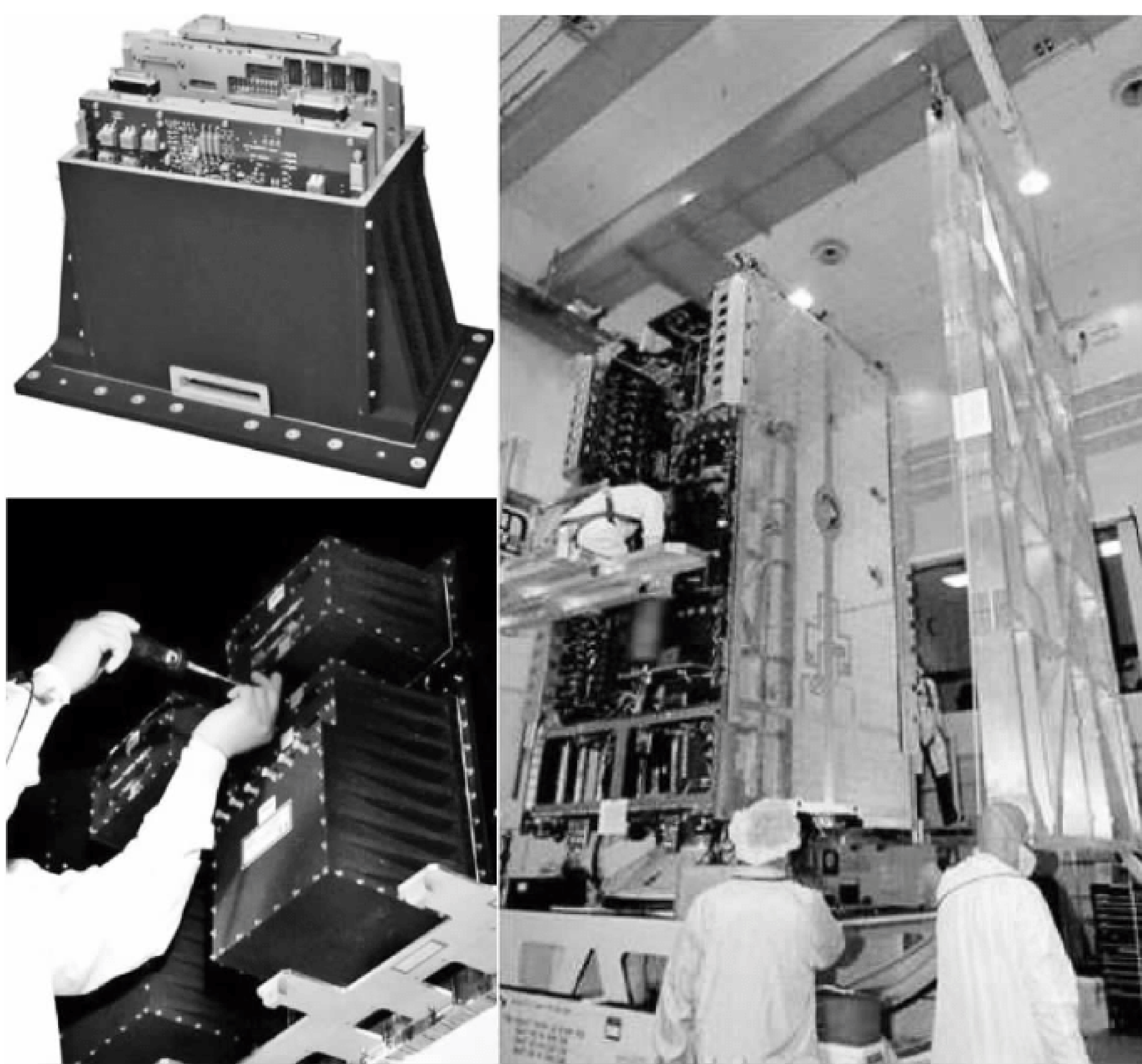


图 6-14　Cisco 太空路由器外形与安装到卫星的工作照片

的综述。深空通信网络协议体系研究主要包括三种思路：空间 IP 协议体系、CCSDS 协议体系和 DTN 协议体系。

1）空间 IP 协议体系

Internet 的成功发展促使人们思考：能不能直接在深空通信网络中应用成熟的 IP 协议与技术。2001 年，美国哥达德航天中心开展了名为 OMNI(Operating Mission as Nodes on the Internet)的研究项目。该项目主要研究利用 IP 协议实现空间通信的可行性，并进行“航天飞机上的通信与导航演示验证(CANDOS)”实验。在空间网络中应用 IP 协议虽然可以基本满足地面与近地轨道航天器间的数据传输，但 TCP 是基于端到端连接与重传的协议，前提是传播延迟很小，这个矛盾在深空通信中被暴露出来。同时，Internet 的路由协议不适用于深空通信网络的结构模型。

2）CCSDS 协议体系

1982 年，美国成立了空间数据系统咨询委员会(CCSDS)，并发布了用于星-地链路与星间链路的从物理层到应用层的一系列建议。CCSDS 针对空间环境的特点，对地面网络的 TCP/IP 标准进行了改进，开发了一套覆盖网络层到应用层的空间通信协议规范(SCPS)，以及 CCSDS 文件传输协议 CFDP。SCPS 并未提出适合深空通信的路由算法，数据传输仍然是建立在先连接后传送数据的端-端模式，CFDP 仅限定于文件传输应用，缺乏更完善的应用服务。由于没有从根本上解决 IP 协议体系固有的缺陷，因此仍然面临着很多需要解决的问题。

3）DTN 协议体系

容迟网络研究小组(DTNRG)提出了一种基于容迟网络 DTN 的协议体系。为解决深空环境下的可靠数据传输问题，JPL 于 2002 年 12 月提交了一份支持 DTN 网络的协议草案，命名为“Licklider”传输协议(LTP)，以便替代 TCP/IP。

目前，深空环境中 DTN 应用研究主要包括路由算法、流量和拥塞控制、安全认证、QoS

与可靠性等方面。路由算法一直是DTN研究中最活跃的领域之一。动态路由算法是地面受限网络DTN应用研究的热点问题，虽然有大量的路由算法研究成果，但很少能够在实际环境中得到验证。DTN路由在异构网络互联的路由策略、链路容量、节点缓存空间及其管理、节点设备处理能力等方面问题仍然没有很好地解决。由于当前深空通信的节点少，拓扑结构简单，似乎路由问题比较简单，但是深空环境链路状态变化大，传输延迟无法控制，因此随着深空任务复杂性不断提高，特定环境的深空网络动态路由问题的研究仍然应该受到足够的重视。DTN网络无法把用户认证、数据传输可靠性工作全部交给接收端的应用层去处理，缺少端-端的控制链路，可靠性实现比较复杂。同时，由于DTN网络资源受限，恶意路由发送伪造的数据束，或者数据束经过未经认证的节点接收、复制和转发，都会对网络安全构成很大的威胁。因此，DTN的流量和拥塞控制、安全认证、QoS与可靠性问题都有待研究。

第三部分　习题参考答案

1. （1）源端口号为1586，目的端口号为53
   （2）用户数据长度为20B
   （3）报文由客户端进程发出
   （4）访问域名解析DNS服务器
2. （1）源端口号为：1330
   （2）目的端口号为：23
   （3）序号为：1
   （4）确认值为：85
   （5）头部长度值为：5
   （6）TELNET
   （7）窗口大小为2047
3. 主机A只能够再发送1000B
4. 确认序号为500
5. 最大吞吐率 为26.214Mbps，信道利用率约为2.62%
6. TCP连接所能够达到的最大传输速率是61.2kbps
7. 新的估计往返延时值为RTT_1=34.1(ms)、RTT_2=33.9(ms)、RTT_3=32.9(ms)。
8. 第1个往返到第15个往返的cwnd值分别为2、4、8、9、10、11、12、1、2、4、6、7、8、9、10
9. 拥塞窗口为9kB
10. ① 10021
    ② 10021
    ③ 25610
    ④ 60036
    ⑤ 16956
    ⑥ 60036
    ⑦ 16956
    ⑧ 60037

第7章 应用层

第一部分　学习目的、要求与知识点结构

1. 学习目的

本章在介绍Internet应用技术发展三个阶段、网络应用与应用层协议分类，以及C/S与P2P模式比较的基础上，将对域名系统DNS、远程登录服务与TELNET协议、电子邮件服务与SMTP、Web与基于Web的网络应用、动态主机配置协议DHCP、网络管理与SNMP进行系统的分析，并通过对典型的应用层协议——FTP执行过程的解析，深入讨论网络应用系统与应用层协议的设计与实现方法。

2. 学习要求

（1）了解：Internet应用的发展与应用层协议的分类。

（2）掌握：Client/Server与P2P模式的特点。

（3）掌握：DNS、DHCP的基本工作原理。

（4）掌握：SMTP、FTP与TELNET等协议的基本工作原理。

（5）掌握：Web与搜索引擎的基本工作原理。

（6）掌握：SIP基本工作原理。

（7）掌握：网络管理SNMP的基本工作原理。

（8）掌握：应用层协议的分析方法。

3. 本章知识点的组织与结构

本章知识点的组织与结构如图7-1所示。

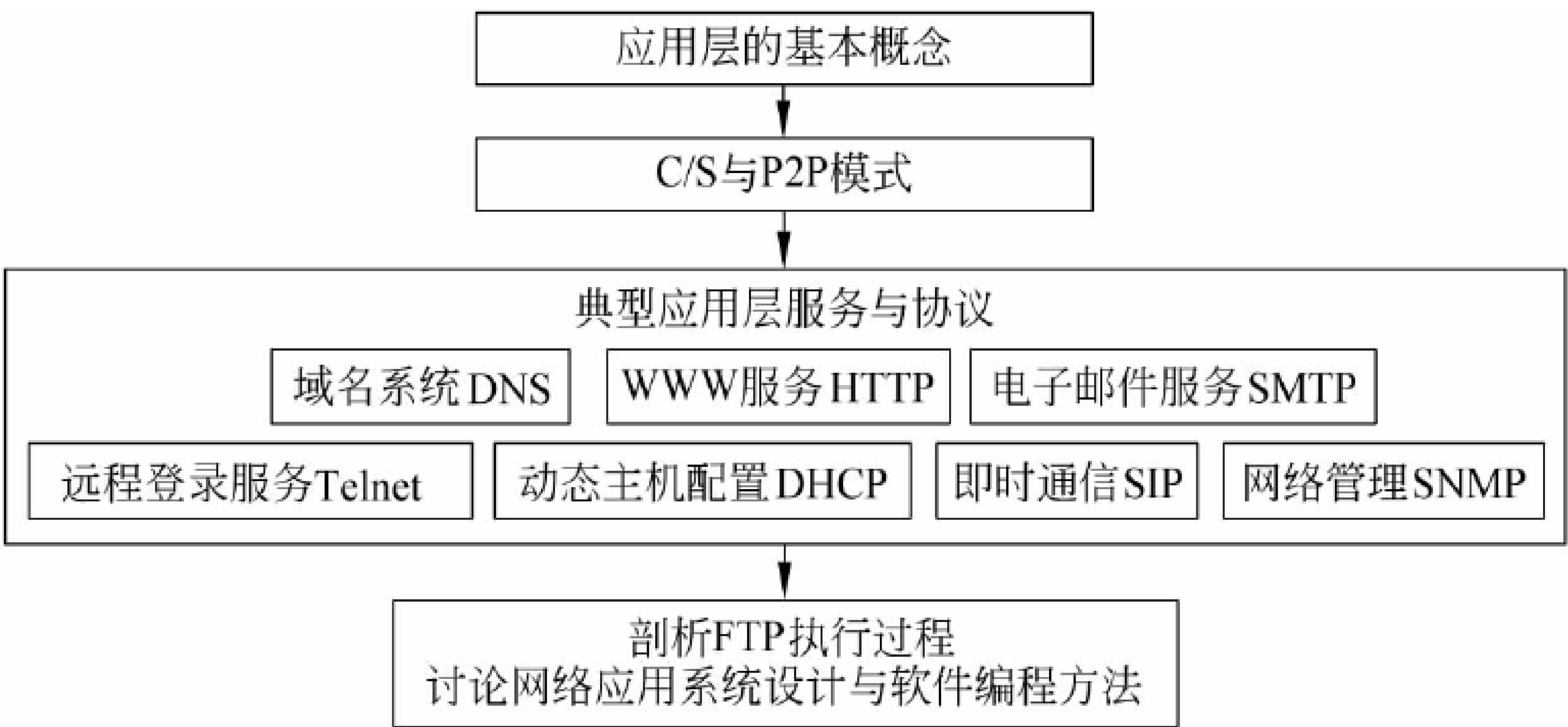


图7-1　第7章知识点的组织与结构

第二部分　教学内容问答

问题 7-1：如何认识 Internet 应用发展不同阶段的特点？

根据作者对 Internet 应用技术发展过程的分析，Internet 的发展大致可以分为三个阶段。图 7-2 给出了 Internet 应用的发展趋势示意图。

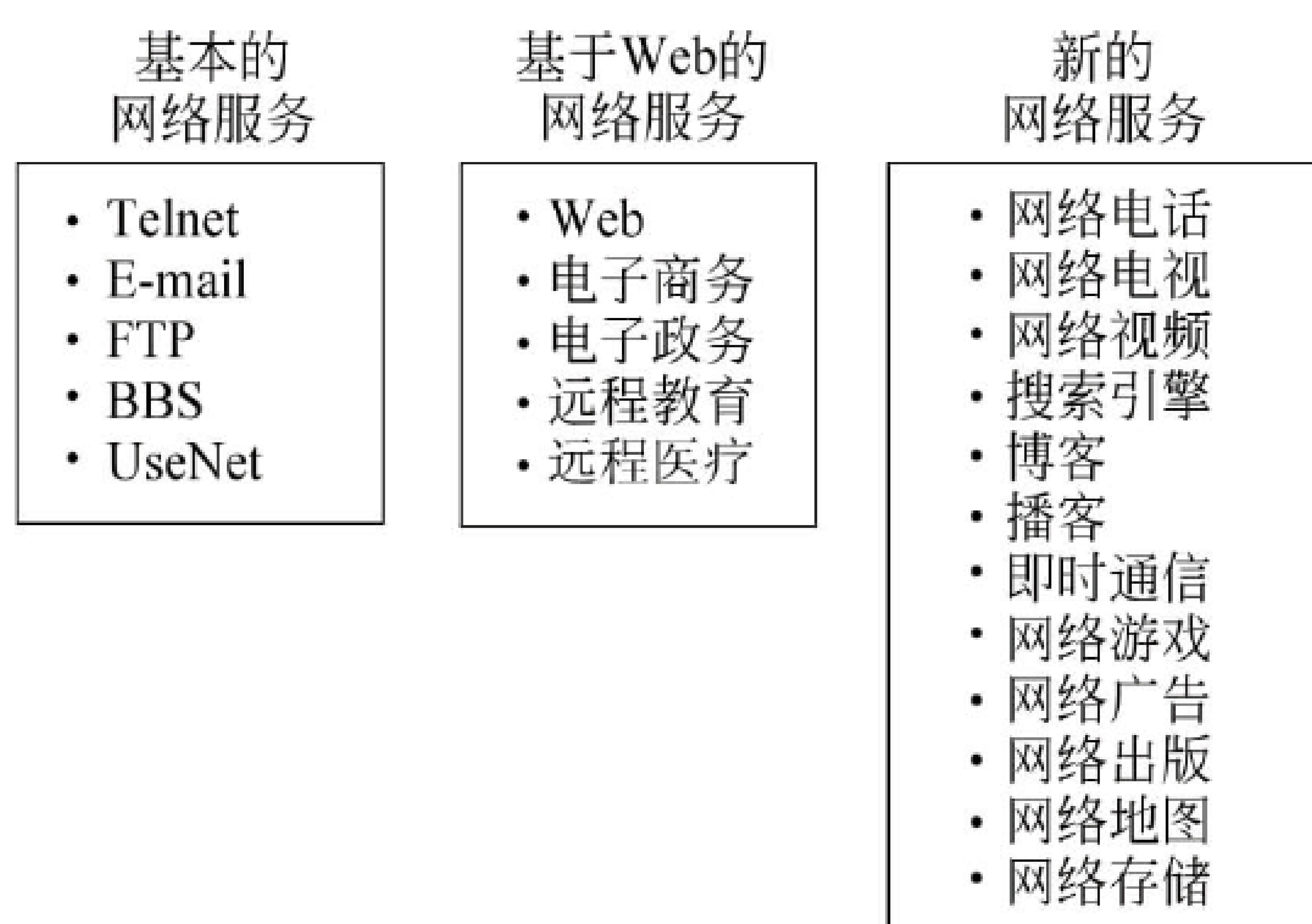


图 7-2　Internet 应用的发展趋势

我们可以根据每一个阶段最具代表性技术的特征，总结出每一个阶段具有以下的主要特征。

1. 第一阶段

第一阶段 Internet 应用的主要特征是：提供远程登录、电子邮件、文件传输、电子公告牌与网络新闻组等基本的网络服务功能。

(1) 远程登录(Telnet)服务实现终端远程登录服务功能。

(2) 电子邮件(E-mail)服务实现电子邮件服务功能。

(3) 文件传输(FTP)服务实现交互式文件传输服务功能。

(4) 电子公告牌(BBS)服务实现网络人与人之间交流信息的服务功能。

(5) 网络新闻组(UseNet)服务实现人们对所关心的问题开展专题讨论的服务功能。

2. 第二阶段

第二阶段 Internet 应用的主要特征是：Web 技术的出现，以及基于 Web 技术的电子政务、电子商务、远程医疗与远程教育应用，以及搜索引擎技术的发展。

3. 第三阶段

第三阶段 Internet 应用的主要特征是：P2P 网络应用扩大了信息共享的模式，无线网络应用扩大了网络覆盖的范围，云计算为网络用户提供了一种新的信息服务模式，物联网扩大了网络技术的应用领域。在云计算环境中，客户可以使用个人计算机、PDA、智能手机、iPad 移动数字终端或家用电器端等“瘦”端系统的设备，随时随地地访问能够提供巨大计算和存储能力的“云服务器”，而不需要知道这些服务器放在什么地方，是什么型号的计算机，使用的是什么样的操作系统和 CPU。

问题 7-2：传输层与应用层都讨论 Client/Server 模式，两者的区别是什么？

客户/服务器(Client/Server，C/S)的概念可以从应用层与传输层两个角度去认识。在

应用程序体系结构的分类中使用的客户/服务器(C/S)的术语对于理解系统结构是有利的,但是容易与应用进程间相互作用模式产生混淆。讨论这个问题需要注意以下几个问题。

(1) 从传输层角度,网络的每项服务都是对应一个"服务"进程。进程通信的实质是实现进程之间的相互作用。网络环境中的进程通信要解决的一个重要问题是确定进程间的相互作用模式。在 TCP/IP 体系中,进程之间相互作用采用的是客户/服务器(C/S)模式。

在客户/服务器(C/S)模式中,客户与服务器分别表示相互通信的两个应用程序进程。客户向服务器发出服务请求,服务器响应客户的请求,提供客户所需要的网络服务。发起本次进程通信、请求服务的本地计算机的进程叫作客户进程,远程计算机提供服务的进程叫作服务器进程。图 7-3 给出了进程通信中的客户/服务器模式。在一次通信过程中,如果主机 A 首先发起一次进程通信,那么主机 A 的进程为客户(Client)进程,而响应的主机 B 的进程为服务器(Server)进程。如果是主机 B 发起的一次进程通信,那么主机 B 的进程为客户进程,而响应的主机 A 的进程为服务器进程。

(2) 我们在讨论应用程序体系结构设计时,是从应用层去分析客户与服务器的概念。例如,用户通过浏览器程序去访问大学网站的 Web 服务器,那么运行浏览器程序的计算机就是客户,而提供 Web 服务的计算机就是服务器。这是从应用层服务的请求服务与提供服务的角度看待客户/服务器(C/S)术语的含义。

(3) 需要注意以下几个问题。

① 使用计算机的人是计算机的"用户(User)",而不是"客户(Client)"。在描述进程间相互作用的客户/服务器模式中,客户(Client)与服务器(Server)分别表示相互通信的两个端系统设备的应用程序进程。在很多文献中,人们也将 Client 译为"客户"或"客户机"。

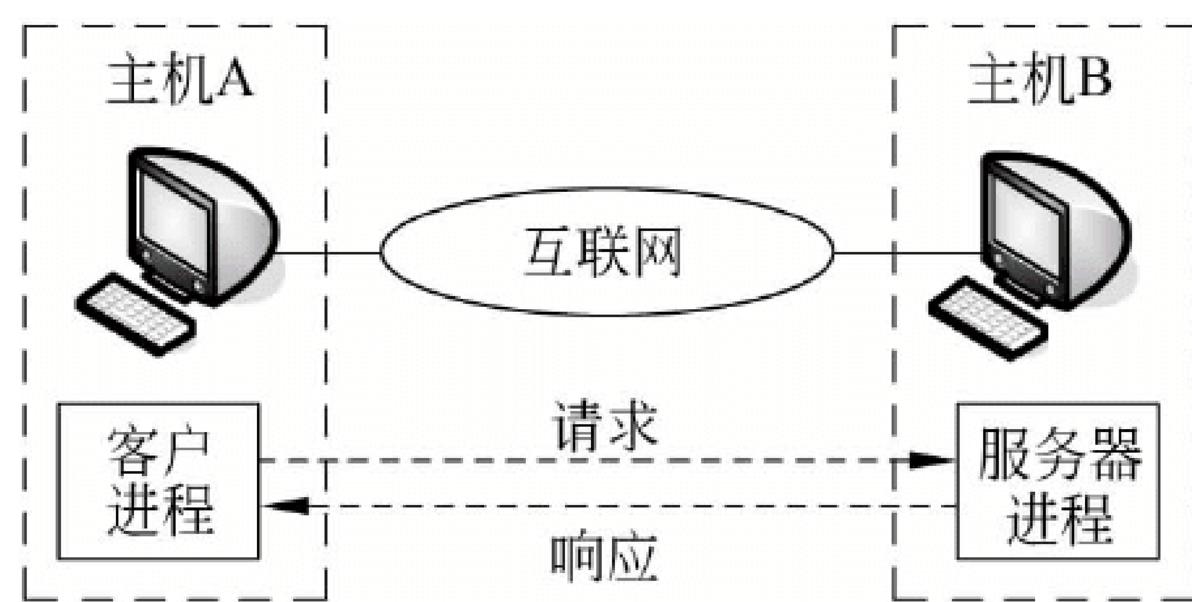


图 7-3 进程通信中的客户/服务器模式

② 在讨论 E-mail、FTP 与 Web 服务时,似乎应用层与传输层的客户与服务器、请求服务与提供服务的角色是一致的,但在 P2P 应用中已经不存在固定的客户与服务器的关系,因此在 P2P 应用中客户与服务器的概念更准确地表现在进程通信的层面。

③ 在传统的互联网中,信息资源的共享是以服务器为中心的 C/S 工作模式。以 Web 服务器为例,Web 服务器是运行 Web 服务器程序、计算能力与存储能力强的计算机,所有 Web 页都存储在 Web 服务器中。服务器可以为很多 Web 浏览器客户提供服务。但是,Web 浏览器之间不能直接通信。显然,在传统互联网信息资源的共享关系中,服务提供者与服务使用者之间的界限是清晰的。

问题 7-3:网络体系结构与应用程序体系结构是什么关系?

为了回答这个问题,需要注意以下几点。

(1) 应用程序体系结构是在网络体系结构的基本框架之下提出的。提出应用程序体系

结构的出发点是为网络应用系统设计与网络软件研发人员提供一种将复杂问题简化的思想方法。

（2）在实际开展一项 Internet 应用系统设计与研发任务的时候，设计者面对的不会只是单一的广域网或局域网环境，而将是多个由路由器互联起来的局域网、城域网与广域网构成的、复杂的 Internet 环境。当复杂的互联网抽象为边缘部分与核心交换部分之后，网络应用系统设计工程师在设计一种新的网络应用时，只需要考虑如何利用核心交换部分能够提供的服务，不涉及核心交换部分的路由器、交换机等低层设备或通信协议软件的编程问题。他的注意力可以集中到运行在多个端系统之上应用程序体系结构的设计与应用软件编程上，这就使得网络应用系统的设计开发过程变得比较容易和规范。

这一点也正体现了网络分层结构的基本思想，也反映出网络技术的成熟。

问题 7-4：为什么说 P2P 网络是在 IP 网络上构建的一种逻辑的覆盖网？

P2P 网络并不是一个新的网络结构，而是一种新的网络应用模式。构成 P2P 网络的结点通常已是 Internet 的结点，它们不依赖于网络服务器，在 P2P 应用软件的支持下以对等方式共享资源与服务，在 IP 网络上形成一个逻辑的网络。这就像在一所大学里，学生在系、学院、学校等各级组织的管理下开展教学和课外活动，同时学校也允许学生自己组织社团，例如计算机兴趣小组、电子俱乐部、博士论坛，开展更加适合不同兴趣与爱好的同学的课外活动。因此，P2P 网络是在 IP 网络上构建的一种逻辑的覆盖网。

问题 7-5：P2P 网络是在什么样的背景下发展起来的？

讨论 P2P 网络发展的背景时，需要注意以下几个问题。

1. 从思维模式变化的角度

网络操作系统设计思想的基础是网络用户资源共享模式。对比网络操作系统的发展过程就会发现，网络操作系统经历了“对等-不对等”发展过程，它为目前网络资源共享的 P2P 技术发展奠定了基础。在 20 世纪 80 年代初出现的很多网络操作系统实际上采取“对等结构”。对等结构网络操作系统的特点是：网络中所有节点安装的网络软件相同，每个节点从资源共享的关系上是平等的。联网的每台主机既是网络服务的提供者，也是网络服务的使用者。联网的主机前台为本地用户提供服务，后台为网络中其他用户提供服务。

2. 从网络资源不断丰富的角度

当联网计算机资源，尤其是硬件资源增强之后，网络操作系统的设计也从“对等结构”发展为“非对等结构”。当联网计算机的硬件资源增强之后，人们可以选择硬件配置好、运算能力与存储能力强的高档个人计算机作为网络服务器，为硬件资源比较差的客户提供服务。非对等结构网络操作系统分为协同操作的两个部分，一部分运行在网络服务器上，另一部分运行在网络客户上。服务器集中管理网络中的共享资源。这些共享的资源主要包括：硬件（存储空间、打印机、通信网关等）、软件与数据。运行在服务器上网络操作系统软件的功能与性能，直接决定网络服务功能的类型、系统性能与安全性。

3. 从用户应用模式变化的角度

在不同技术发展阶段，人们对网络应用关注的重点也不同。初期阶段重点是在共享网络硬件上。中期阶段重点是在共享软件和数据上。到成熟阶段，重点应该转移到共享信息资源上。这正反映出用户希望自己在 Internet 中扮演角色的转变。用户开始不满足只作为

信息资源的享受者,希望能同时扮演信息享受者和信息提供者的双重身份,这也正反映出用户网络应用水平的提高和网络作用的深化。在计算机硬件配置提高,网络应用水平提高,网络信息资源积累与存储格局变化的基础上,必将导致网络资源共享模式的变化,在这样的技术发展背景下出现的P2P网络的发展也就显得很自然。

问题7-6:如何认识域名系统、域名与域的关系?

为了理解这个问题,需要注意以下几点。

1. 域名系统

域名系统是Internet使用的命名系统。实际上,人们将主机的名字叫作域名,其原因是Internet使用的命名系统定义了很多的域。主机要按照它所属的域来命名,因此主机名又叫作域名。

2. 域

域名空间结构被分成两百多个顶级域TLD,常用的通用域有:.com(商业)、.edu(教育性机构)、.gov(政府)、.net(网络服务供应商)、.org(非营利性机构)、.int(国际性组织)、.mil(军事组织),以及国家级或地区域名,每个域自己控制如何分配它下面的域。

3. 域名

域名是Internet中主机按照一定的规则,用自然语言(英文缩写的域名或中文域名)表示的名字,它与签订的IP地址项对应。

问题7-7:域名、端口号、IP地址、MAC地址是什么关系?

为了回答这个问题,需要注意以下几点。

(1) 域名:域名是应用层使用的主机名字。例如以www.netlab.cs.nanaki.edu.cn为例,人们可以很容易将它理解为"中国-教育机构-南开大学-计算机系-网络实验室-Web服务器"。

(2) 端口号:端口号是传输层的进程通信中用于标识进程的号码。例如,如果希望访问服务器www.netlab.cs.nanaki.edu.cn,需要使用Web服务器进程的熟知端口号80。

(3) IP地址:IP地址是网络层IP协议使用的地址。例如,对应www.netlab.cs.nanaki.edu.cn的IP地址是201.1.2.16。

(4) MAC地址:MAC地址是MAC层帧传输过程中使用的地址。例如,与IP地址201.1.2.16对应的网卡的MAC地址为07-00-1A-20-00-28。

总结以上的讨论可以看出:如果用户在一台计算机上通过浏览器访问一台Web服务器时,需要使用客户计算机与服务器计算机的域名、端口号、IP地址、MAC地址来唯一地标识主机,寻址、路由、传输,实现网络环境中的分布式进程通信,完成Internet的访问过程。图7-4给出了域名、端口号、IP地址与MAC地址的关系。

问题7-8:区、域与域名服务器是什么关系?

理解域名服务器的工作原理,需要注意以下几个问题。

DNS的本质是:提出一种分层次、基于域的命名方案,并且通过一个分布式数据库系统,以及维护与查询机制来实现域名服务功能。DNS必须具备以下三种基本功能。

(1) 域名空间定义:定义一个包括所有可能出现结点的域名空间。

(2) 域名注册:为每台主机分配一个在全网具有唯一性的域名。

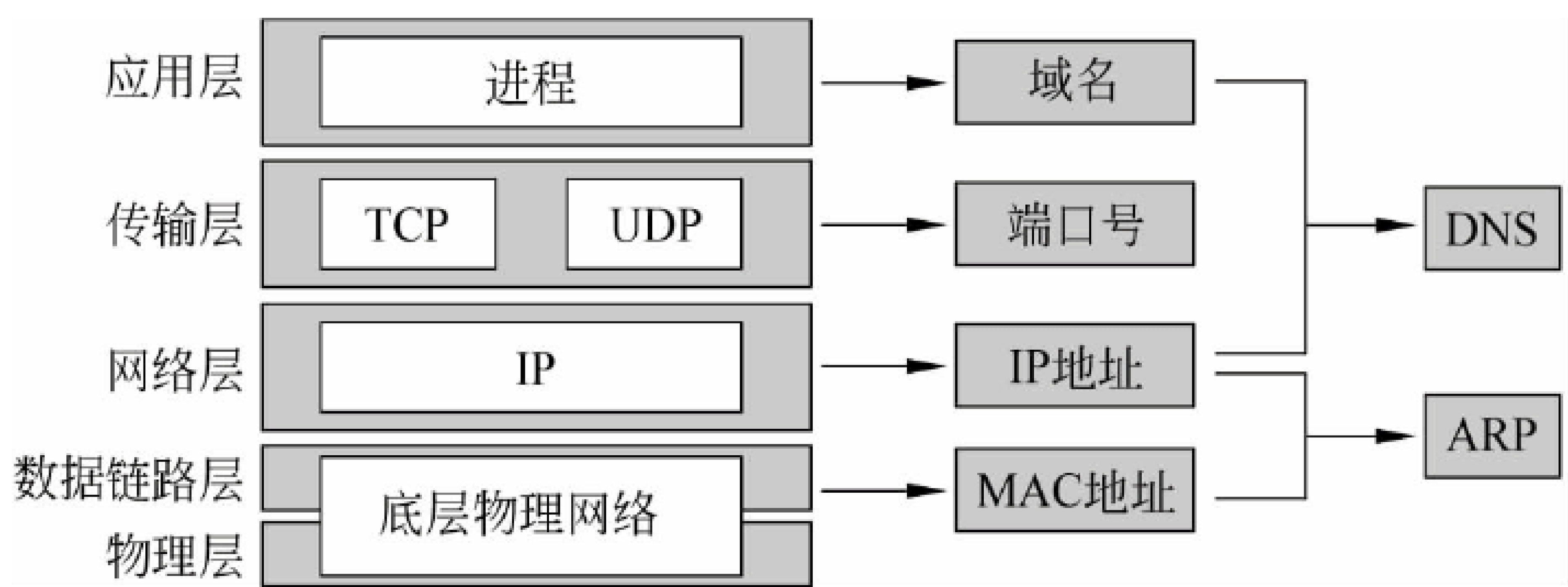


图 7-4　域名、端口号、IP 地址与 MAC 地址的关系

(3) 域名解析：为用户提供一种有效的完成域名与网络 IP 地址转换的机制。

相应的 DNS 包括三个组成部分：域名空间、域名服务器、地址解析程序。

图 7-5 给出我国大学域名管理的一个例子。南开大学作为一个独立的行政单位，它被管理中国教育科研网(CERNET)网络中心授权管理"nankai. edu. cn"的域，因此由南开大学校园网中心管理"nankai. edu. cn"域。设置管理"nankai. edu. cn"域的域名服务器可以用一种最简单的办法，那就是只设置一个域名服务器，管理所有南开大学内部的域名。但是，一个单位规模太大了，这种集中管理的方法带来的问题是域名系统运行效率低，不能够满足用户服务质量要求。最有效的方法如下。

(1) 根据需要将一个"域"划分成不重叠的多个"区"。

(2) 每个"区"设置相应的权限域名服务器，用来保存该区内所有主机的域名与 IP 地址的映射关系数据。"区"是域名服务器管辖的范围。

(3) 各个"区"的域名服务器都相互连接，构成支持整个"域"的域名服务器体系。

图 7-5(a)表示一个域没有划分区的情况，那么区就等于域，只要设置一个域名服务器就可以管理整个校园网的域名。图 7-5(b)表示一个域没有划分两个区的情况，那么这两个区 nankai. edu. cn 与 it. nankai. edu. cn 都属于 nankai. edu. cn 的域。图 7-5(c)表示在两个区分别设置具有相应权限域名服务器的结构。一个域名服务器有权管管辖的范围叫作"区"，它是"域"的一个子集。

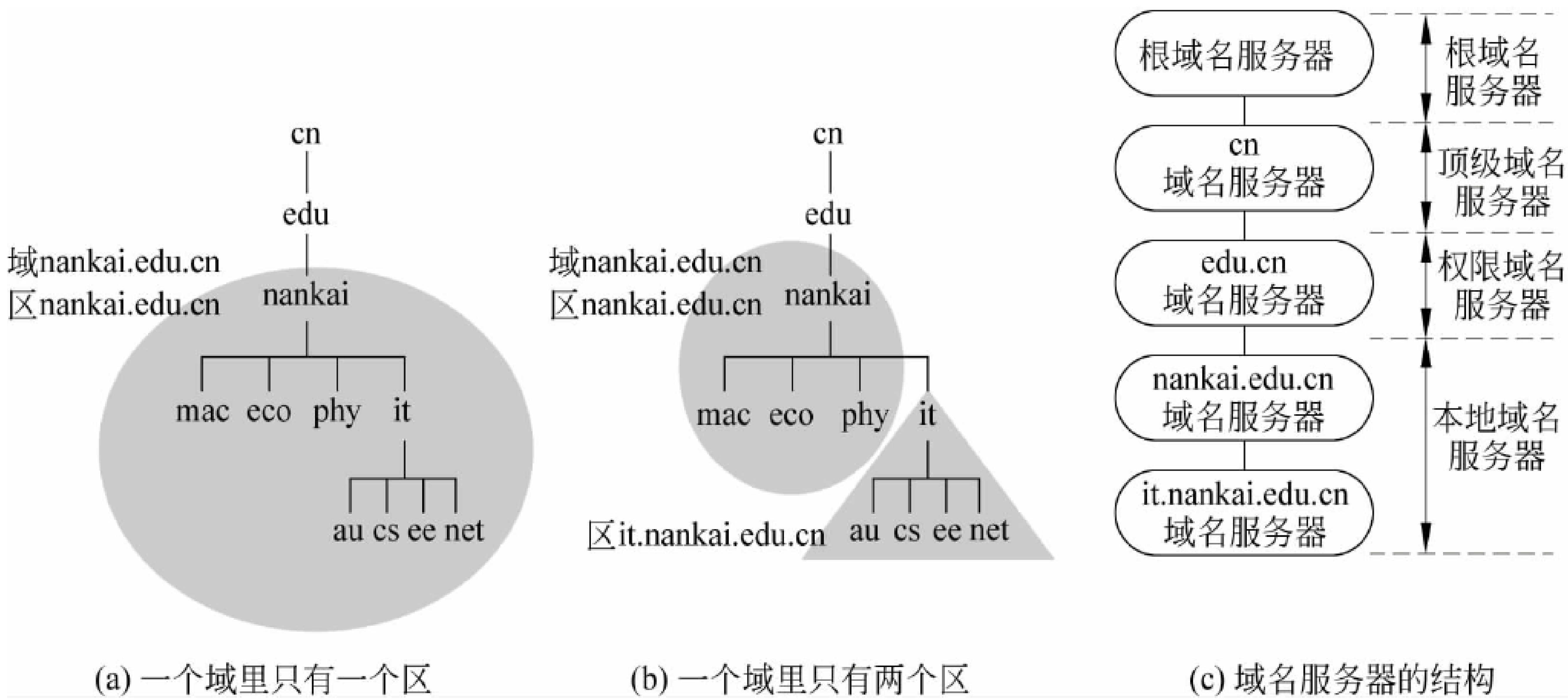


图 7-5　域、区与域名服务器

问题 7-9：域名服务器有几种类型?

(1) 支持 Internet 运行的域名服务器是按层次来设置的,每一个域名服务器都只对域名空间中的一部分进行管辖,由多个层次结构的域名服务器系统覆盖整个域名空间。

(2) 根据域名服务器所处的位置和所起的作用,域名服务器可以分为以下 4 种类型。

① 根域名服务器。

根域名服务器对于 DNS 系统的整体运行具有极为重要的作用。任何原因造成根域名服务器停止运转,都会导致整个 DNS 系统的崩溃。出于安全的原因,目前存在的 13 个 DNS 根域名服务器,其专用域为 root-server. net。大多数根域名服务器是由一个服务器集群组成。有些根域名服务器是由分布在不同地理位置的多台镜像 DNS 服务器组成,例如,根域名服务器 f. root-server. net 就是由分布在四十多个地方的几十台镜像 DNS 服务器组成。有关最新的根域名服务器列表可以从 ftp://ftp. rs. internic. net/domain/named. root 中获取。

② 顶级域名服务器。

顶级域名服务器负责管理在该顶级域名注册的所有二级域名。例如,在中国互联网信息中心 CNNIC 管理所有在“. cn”之下注册的通用域名与行政区域域名。

③ 权限域名服务器。

权限域名服务器负责经过授权的一个区的域名管理。

④ 本地域名服务器。

本地域名服务器也叫作默认域名服务器。一个 ISP、一所大学甚至一个系都可能有一个或多个本地域名服务器。

为了保证域名服务器系统的可靠性,域名服务器一般需要将域名数据复制到几个域名服务器上,其中一个为主域名服务器,其他的是从域名服务器。主域名服务器定期将数据复制到从域名服务器;当主域名服务器出现故障时,从域名服务器继续执行域名解析的任务。

问题 7-10：ARP 与 DNS 是什么关系?

图 7-6 给出了 ARP 与 DNS 关系示意图。从图中可以看出:

(1) ARP 是用于解析网络层 IP 地址与数据链路层 MAC 地址的对应关系。

(2) DNS 协议是用于解析应用层主机域名与网络层 IP 地址的对应关系。

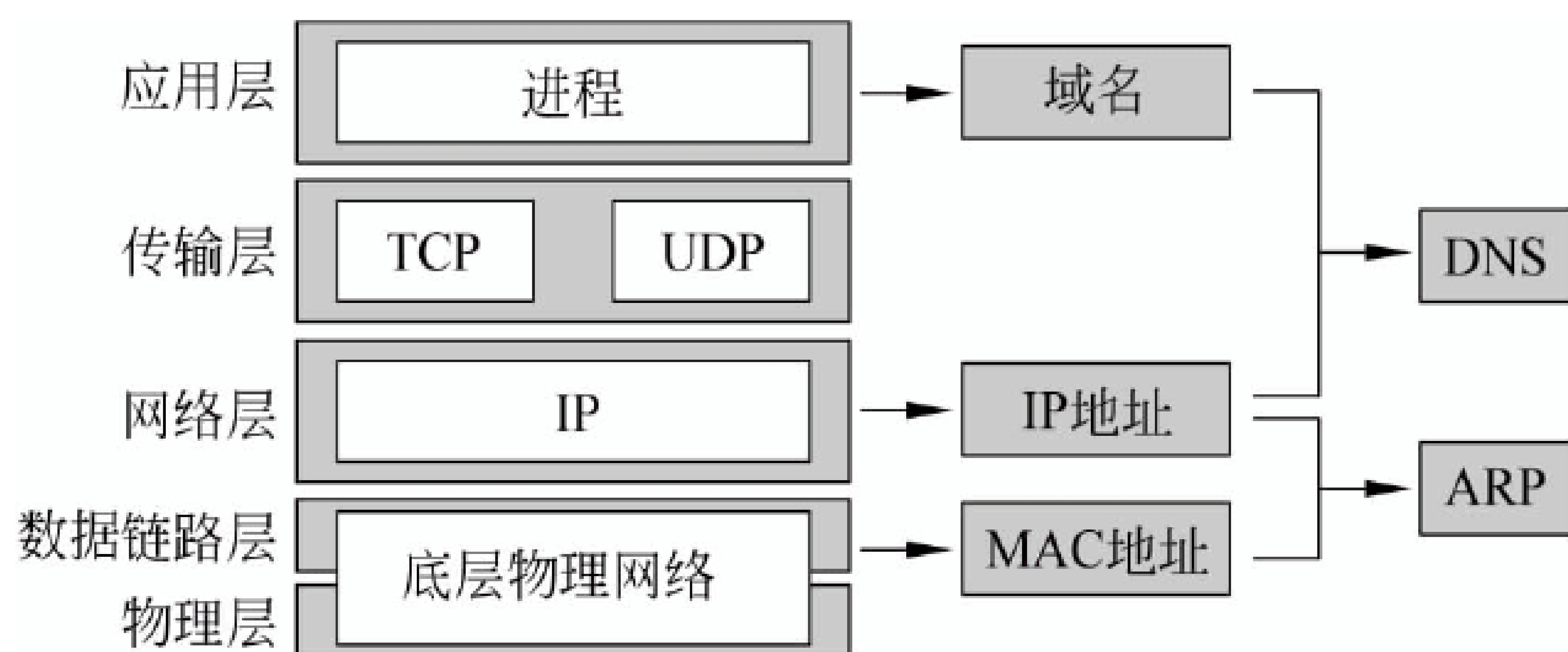


图 7-6　ARP 与 DNS 的关系示意图

问题 7-11：TELNET 协议是在什么样的背景下产生的?

TELNET 协议出现在 20 世纪 60 年代后期,那时个人计算机(PC)还没有出现。当时

人们在使用大型计算机时，必须通过直接连接到主机的某一个终端，在使用用户名与密码登录成为合法用户之后，才能将软件与数据输入到主机，完成科学计算的任务。当用户需要使用多台计算机共同完成一个较大的计算任务时，需要调用远程计算机与本地计算机协同工作。当这些大型计算机互联之后，就需要解决一个问题，那就是不同型号计算机之间的差异性问题。

不同型号计算机系统的差异性主要表现在硬件、软件与数据格式上。最基本的问题是：不同计算机系统在对终端键盘输入命令的解释就不同。例如，有的系统用 return 或 enter 作为行结束标志，有的系统用 ASCII 字符的 CR，而有的系统用 ASCII 字符的 LF。键盘定义的差异给远程登录带来很多问题。在中断一个程序时，有些系统使用“^C”，而另一些系统使用 Esc 键。发现这个问题之后，各个厂商都分别研究如何解决互操作性的方法，例如，Sun 公司制定远程登录协议 rlogin，但是该协议是专为 BSD UNIX 系统开发的，它只适用于 UNIX 系统，并不能很好地解决不同类型计算机之间的互操作性问题。为了解决异构计算机系统互联中存在的问题，人们开始研究 TELNET 协议。

问题 7-12：为什么 TELNET 协议又称为网络虚拟终端协议？

TELNET 协议引入网络虚拟终端(Network Virtual Terminal，NVT)的概念，它提供一种专门的键盘定义，用来屏蔽不同计算机系统对键盘输入的差异性，同时定义客户与远程服务器之间的交互过程。TELNET 协议的优点是能解决不同类型的计算机系统之间的互操作问题。远程登录服务是指用户使用 TELNET 命令，使自己的计算机暂时成为远程计算机的一个仿真终端的过程。当用户成功实现远程登录后，用户计算机就可以像一台与远程计算机直接相连的本地终端一样工作。因此，TELNET 协议又称为网络虚拟终端协议或远程终端协议。

问题 7-13：TELNET 协议如何实现异构计算机系统之间的相互访问？

图 7-7 给出了 TELNET 的工作原理示意图。用户的实终端采用用户终端的格式与本地 TELNET 客户通信；远程计算机采用主机系统格式与 TELNET 服务器通信。在 TELNET 客户进程与 TELNET 服务器进程之间，通过网络虚拟终端(Network Virtual Terminal，NVT)标准来进行通信。NVT 是一种统一的数据表示方式，以保证不同硬件、软件与数据格式的终端与主机之间通信的兼容性。

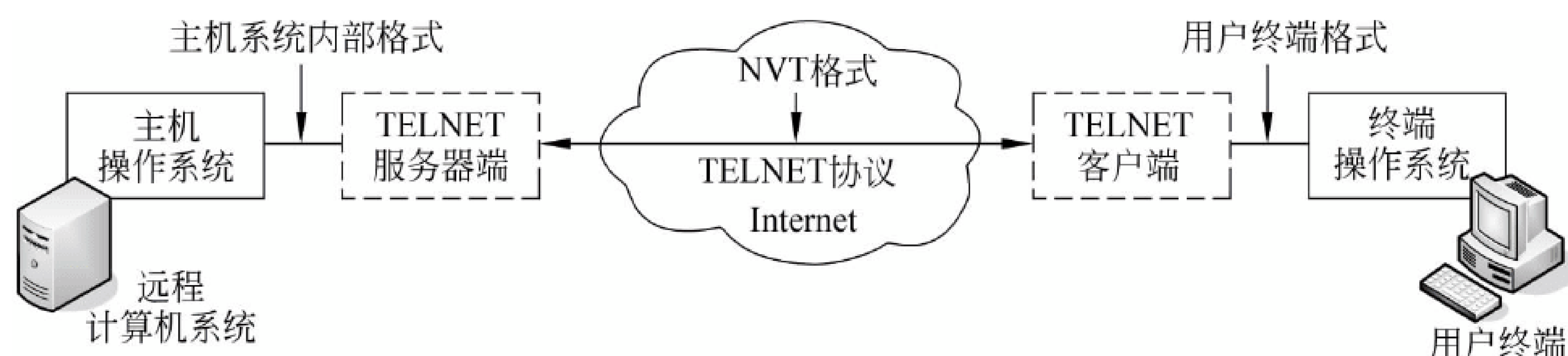


图 7-7　TELNET 的工作原理示意图

TELNET 客户端将用户终端发出的本地数据格式转换成标准的 NVT 格式，再通过网络传输到 TELNET 服务器端。TELNET 服务器将接收到的 NVT 格式数据转换成主机内部数据格式，再传输给主机。Internet 上传输的数据都是 NVT 格式。在引入网络虚拟终端概念之后，不同的用户终端与服务器进程将与各种不同的本地终端格式无关。TELNET 客

户与服务器进程完成用户终端格式、主机系统内部格式与标准NVT格式之间的转换。

问题7-14：电子邮件系统运行过程中涉及哪几种协议？

回答这个问题需要注意以下几点。

1. Internet电子邮件系统设计的原则

(1) 所有邮件都使用标准的地址格式，并且每个邮箱地址是唯一的。

(2) 所有邮件报文都使用统一的报文格式，从而保证不同系统之间可以交换邮件。

(3) 邮件的发送方和接收方都使用统一的邮件传输协议对邮件进行传送。

2. 电子邮件系统运行过程中涉及的协议

电子邮件系统分为两个部分：邮件服务器端与邮件客户端。

(1) 在邮件服务器端，包括用来发送邮件的SMTP服务器，用来接收邮件的POP3服务器或IMAP服务器，以及用来存储电子邮件的电子邮箱；在邮件客户端，包括用来发送邮件的SMTP代理，用来接收邮件的POP3代理，以及为用户提供管理界面的用户接口程序。

(2) 邮件客户端使用简单邮件传输协议(Simple Mail Transfer Protocol，SMTP)向邮件服务器发送邮件；邮件客户端使用邮局协议(Post Office Protocol，POP)的第3版POP3协议或交互式邮件存取协议(Interactive Mail Access Protocol，IMAP)，从邮件服务器中接收邮件。至于使用哪种协议接收邮件，取决于邮件服务器与邮件客户端支持的协议类型，一般的邮件服务器与客户端应用程序支持POP3协议。

问题7-15：MIME邮件传输协议怎样扩展SMTP功能？

回答这个问题需要注意以下几点。

(1) 最初出现的描述SMTP的RFC文档出现在1982年。由于受到当时网络带宽的限制，只允许Internet电子邮件协议就像它的名称一样，只能是一个简单的邮件传输协议。SMTP的局限性表现在只能发送NVT-7 ASCII码格式的报文，不支持那些不使用NVT ASCII码格式的语种，例如中文、法文、德文、俄文、希伯来文等。同时，它也不支持语音、视频的数据。

(2) 通用协议Internet邮件扩展(Multipurpose Internet Mail Extension，MIME)是一种辅助性的协议，它本身不是一个邮件传输协议，只是对SMTP的补充，并不能替代SMTP。MIME的功能是允许非7b NVT ASCII码格式的数据通过SMTP传输。

(3) 这里涉及网络虚拟终端(NVT)的概念。NVT的概念是在Internet的远程登录服务TELNET中提出的。TELNET允许用户使用本地终端访问远程计算机系统的应用程序。但是，在网络环境中，本地终端所使用的操作系统与远程计算机使用的操作系统可能是异构的。因此，本地终端所使用的字符，远程计算机操作系统可能不能识别或被错误地解释。为了解决这个问题，远程登录服务TELNET制定了一组网络虚拟终端NVT的通用字符集。本地终端输入的字符首先由本地TELNET转化成NVT格式，通过网络传输的是NVT格式的字符，远程TELNET再将NVT格式转化成远程计算机操作系统能识别的字符格式。

问题7-16：什么是Web服务的基本和核心的协议？

Web服务的核心技术是：超文本传送协议(HTTP)、超文本标记语言(HTML)、超链接(Hyperlink)。最基本和核心的是HTTP。

超文本传输协议 HTTP 的第一个版本是 HTTP0.9，它实现了最基本的文本信息的传输功能。1996 年，RFC1945 给出的 HTTP1.0 在文本传输的基础上，增加了语音、图形与视频的多媒体信息传输的功能。HTTP0.9 与 HTTP1.0 只支持非持续连接方式，因此 Web 系统工作效率比较低。1997 年，RFC2068 给出的 HTTP1.1 增加了持续连接工作方式，以及高速缓存等功能，使 Web 系统的工作效率有很大提高，应用范围也得到迅速拓宽。

问题 7-17：URL 的作用是什么？

理解 URL 的作用需要注意以下几个问题。

（1）统一资源定位符 URL 是对 Internet 资源的位置和访问方法的一种简洁表示。只要能够对资源定位，计算机系统就可以对资源进行存取、更新、替换和查找等各种操作。这里所说的“资源”是指在 Internet 上可以被访问的任何对象，包括文件目录、文件、文档、图像、声音等，以及电子邮件的地址、USENET 新闻组或 USENET 新闻组中的文档。标准的 URL 由三个部分组成：协议类型、主机名和路径及文件名。

（2）通过使用 URL 机制，用户可以指定要访问什么服务器、哪台服务器、服务器中的哪个文件。

（3）除了通过指定 HTTP 来访问 Web 服务器之外，URL 还可以通过指定其他协议类型来访问其他类型的服务器。例如：

Gopher://gopher.cernet.edu.cn 表示要连接到名为 gopher.cernet.edu.cn 的 Gopher 服务器。

ftp://ftp.pku.edu.cn/pub/dos/readme.txt 表示要通过 FTP 连接来获得一个名为 readme.txt 的文本文件。

File://linux001.nankai.edu.cn/pub/gif/wu.gif 表示要在所连接的主机上获得并显示一个名为 wu.gif 的图形文件。

telnet://cs.nankai.edu.cn 表示远程登录到名为 cs.nankai.edu.cn 的主机。

问题 7-18：如何理解 HTTP 无状态协议的特征？

HTTP 在传输层使用的是 TCP。如果 Web 浏览器想访问一个 Web 服务器，那么作为客户端的 Web 浏览器就需要与 Web 服务器之间建立一个 TCP 连接。一旦 TCP 连接建立之后，客户端的 Web 浏览器进程就可以发送 HTTP 请求报文，接收应答报文。同样，Web 服务器也可以接收 HTTP 请求报文，发送应答报文。由于 TCP 提供的是面向连接的可靠服务，这就意味着 Web 客户进程发送的 HTTP 请求报文可以正确到达服务器端。同时，Web 服务器进程发送的 HTTP 应答报文也可以正确到达客户端。即使报文在传输过程中出现丢失与乱序的问题，也由传输层及一些低层协议去解决，Web 浏览器与 Web 服务器端进程不需要干预。因此，尽管 TCP 是面向连接的，但是 HTTP 是无连接的。

这里需要注意 HTTP 的一个规定，由于 Web 服务器要面对很多浏览器的并发访问，为了提高 Web 服务器对并发访问的处理能力，因此在设计 HTTP 时规定 Web 服务器发送的 HTTP 应答报文和文档时，不保存任何发出请求的 Web 浏览器进程状态信息。这就可能出现一个 Web 浏览器在短短几秒钟之内两次访问同一对象时，服务器进程不会因为已经给它发出过应答报文而不接受第二次服务请求。由于 Web 服务器进程不保存发出请求的 Web 浏览器进程的任何状态信息，因此 HTTP 属于无状态协议。

问题7-19：如何理解HTTP非持续连接与持续连接、非流水线与流水线的特征？

为了理解这个问题，需要注意以下几点。

1. 非持续连接与持续连接

HTTP支持两种连接模式：非持续连接与持续连接。其中，HTTP1.0版协议定义非持续连接，HTTP/1.1默认状态为持续连接。

1）非持续连接

如果一个网页包括一个基本的HTML文件和105个JPEG图像文件，那么就称为这个Web页是由106个对象(Object)组成。对象就是文件，例如HTML文件、JPEG、GIF图像文件、Java程序、语音文件等，它们都可以通过URL来寻址。在非持续连接中，对每次请求/响应都要建立一次TCP连接。如果一个网页包括106个对象，并且都保存在同一个服务器中，非持续连接的缺点是：必须为每个请求对象建立和维护一个新的TCP连接。对于每个这样的连接，客户端与服务器端都需要设定缓冲区及其他变量。服务器在处理大量客户进程端请求时负担很重，因此研究持续连接的HTTP势在必行。

2）持续连接

在持续连接时，服务器在发出响应后保持该TCP连接，在相同的客户进程端与服务器端之间的后续报文都通过该连接传送。如果一个网页包括一个基本的HTLM文件和8个JPEG图像文件，所有请求与应答报文都通过这个连接来传送。这样，一个完整的网页(包括一个基本的HTLM文件和多个图形文件)，可以通过一个持续的TCP连接来传送。同时，一个Web服务器中的多个Web页也可以通过一个持续的TCP连接来传送。服务器进程在接收到客户进程的请求或超时才关闭该连接。

2. 非流水线与流水线方式

持续连接又有两种工作方式：非流水线与流水线方式。

1）非流水线方式

非流水线方式的特点是：客户端只有在接收到前一个响应时才能发出新的请求。这样，客户端在每访问一个对象时要花费一个RTT时间。这时，服务器每发送一个对象之后，要等待下一个请求的到来，连接处于空闲状态，浪费了服务器的资源。

2）流水线方式

流水线方式的特点是：客户端在没有收到前一个响应时就能够发出新的请求。客户端的请求可以像流水线作业一样，连续地发送到服务器端，服务器端可以连续地发送应答报文。使用流水线方式的客户端访问所有对象只需花费一个RTT时间。因此，流水线方式可以减少TCP连接的空闲时间，提高下载Web文档的效率。HTTP1.1默认状态是持续连接的流水线工作方式。

问题7-20：B/S模式与C/S模式到底有哪些区别？

在讨论C/S模式与B/S模式的区别与联系时，需要注意以下几个问题。

(1) C/S模式作为传输层术语与应用层术语时，它的内涵是有很大区别的。

在传输层与应用层我们都会遇到术语“C/S模式”。C/S模式作为传输层与应用层术语时内涵是有很大的区别的。

① 在讨论了传输层的通信进程之间相互作用方式时用到C/S模式的术语。在传输层

发起一次进程通信的一方为客户(Client)进程,响应一方的进程叫作服务器(Server)进程。因此,Client 进程与 Server 进程是两个通信进程相对的关系,不是专指某一台计算机必须是 Server,而某台计算机就只能作 Client。

② 在应用层 C/S(客户/服务器)模式术语的内涵不同。Internet 应用软件分为服务器软件与客户软件,运行服务器软件的计算机称为“Server”,而运行客户软件的计算机称为“Client”。为了提高网络应用程序的服务质量,人们自然会选择高配置、高性能的计算机作为服务器,而任何一部台式计算机、笔记本,甚至是一个智能手机、数字终端设备都可以成为一个客户机。

(2) B/S 模式与 C/S 模式属于应用层网络应用软件开发与应用模式的问题。

浏览/服务器(Browser/Server,B/S)模式是随着 Web 技术的发展,在应用层 C/S 工作模式基础上演变出来的。

在 Web 技术出现之前,早期基于 C/S 工作模式的各种网络应用软件的客户端软件(例如企业管理软件)的开发有两个重要特点。一是需要为不同的应用软件开发不同的服务器程序与用户程序;二是即使是在图形用户界面操作系统 Windows 之上开发的用户界面,一般还是采用滚动条方式,例如 DOS 字符界面,用户程序模块是按树状结构组织的。因此用户使用一种功能,需要从用户界面入口逐层进入,然后再逐层退出,并且每种程序都是专门开发的,它们互不相同,用户使用每种程序时都需要进行专门的培训。

随着 Internet 和 Web 技术的推广,大批 Internet 用户都熟悉 Web 浏览器的图形用户界面,如果仿照 Web 服务器与浏览器技术来开发基于网络的应用软件和各种管理软件,选择服务器、设计网络体系结构,可以让用户在使用一种新的管理软件时就和他用浏览器去访问 Web 页一样方便。因此,基于 B/S 模式开发网络应用软件可以规范企业网络系统结构设计方法,简化程序员的开发与软件升级过程,简化用户培训过程,方便用户使用。

Intranet 就是在 B/S 模式企业网络应用系统出现之后产生的术语。但是需要注意的是,基于 B/S 模式开发的企业网络应用软件并不一定是运行在 Internet 和 Web 体系中,大量基于 B/S 模式的企业网络应用系统是建立在企业专网之中的。当然,在采用必要的安全防范措施之后,用户可以直接通过 Internet,在异地或移动过程中访问内部服务器,处理办公事务。

(3) 需要注意几种提法。

在讨论 C/S 模式与 B/S 模式区别与联系时,有如下几种提法值得商榷。

① 第一种提法:C/S 一般建立在专用的网络上,小范围里的网络环境,局域网之间再通过专门的服务器提供连接和数据交换服务。B/S 建立在广域网之上的,比 C/S 的适应范围更广,一般只要有操作系统和浏览器就行。

这种提法不准确。C/S 模式与 B/S 模式属于不同时期应用层软件编程的两种基本的模式问题,与它们是否只能够用于局域网与广域网无关。很多大型的网络应用软件是按 C/S 模式开发的,但是它覆盖全国范围的相关部门。同时,一个政府内网只在一座办公大楼里使用,它是按 B/S 模式开发的。

② 第二种提法:B/S 建立在广域网之上,对安全的控制能力相对弱,面向的是不可知的用户群。因此 C/S 模式安全,B/S 模式不安全。

这种提法不准确。C/S 模式与 B/S 模式的安全性取决于系统安全性设计水平。C/S

模式的应用软件是专门为一个部门、一种应用设计、开发的,用户使用一般需要培训。但是,人们不熟悉不表示它就是安全的。B/S模式容易使用,但不表示它不安全。

③ 第三种提法:B/S结构管理软件只安装在服务器端上,网络管理人员只需要管理服务器就行了,用户事务在前端实现,所有的客户端只有浏览器,网络管理人员只需要做硬件维护。

这种提法也是不准确的。从网络管理员管理的角度,C/S模式与B/S模式构建的系统没有本质的区别,基本是相同的。

④ 第四种提法:C/S模式是客户机与服务器两级结构,B/S模式是客户机与文件服务器、数据库服务器三级结构;在C/S体系下,数据库不能真正成为公共、专业化的仓库,它受到独立的专门管理。C/S程序一般是典型的中央集权的机械式处理,交互性相对低。

这种提法也是不准确的。无论是C/S模式或B/S模式都可以采用两级结构或三级结构,数据库是否作为公共、专业化的仓库与独立的专门管理,以及交互性问题,这些与C/S模式或B/S模式不存在必然的联系,取决于设计者对总体结构的选择与软件编程思路。

问题7-21:搜索引擎的基本工作原理是什么?

理解搜索引擎的基本工作原理时,需要注意以下几个问题。

1. 搜索引擎的基本工作原理

当用户在使用搜索引擎时,首先要提交一个或多个“关键字”(或检索词),通过浏览器输入搜索引擎的界面。搜索引擎返回与“关键字”相关的信息列表,它通常包括三方面内容:标题、URL、摘要。其中,标题是从网页的<TITLE></TITLE>标签中提取的内容;URL是网页的访问地址;摘要是从网页内容中提取。用户需要浏览这些内容,挑选自己真正需要的内容,然后通过对应的URL访问该网面。由于不同读者对信息的需求相差很大,即使同一读者在不同时间关心的问题也不同,因此搜索引擎不可能理解读者的真正需求,只能争取做到尽可能不漏掉任何有用信息。由于反馈给用户的很长的列表经常使读者感到困惑和无从下手,因此搜索引擎还要将用户“最可能关心的信息”排在列表前面。

搜索引擎技术起源于传统的全文检索理论。全文检索程序通过扫描一篇文章中所有词语,并根据检索词在文章中出现的频率和概率,对所有包含这些检索词的文章进行排序,最终给出可以提供给读者的列表。

2. 搜索引擎的基本结构

基于全文搜索的搜索引擎通常包括4个部分:搜索器、索引器、检索器与用户接口。

1) 搜索器

搜索引擎通过搜索器在Internet上逐个访问Web站点,并建立一个网站的关键字列表。人们将搜索器建立关键字列表的过程称为“爬行”。搜索器要根据一个事先制定的策略确定一个URL列表,而这个列表通常是从以前访问的记录中提取,特别是一些热门站点和包含新信息的站点。搜索器访问每个Web站点后,需要分析与提取新的URL,并将它加入访问列表中。搜索器遍历指定的Web空间,将采集到的网页信息添加到数据库。但是,采集Internet上的所有网页是不可能的。最大的搜索引擎抓取的网页也只可能占40%。在建立初始网页集时,最可能的方法是启动多个搜索器,并行地访问多个Web站点的网页。

实际上,每个搜索器的搜索策略与过程都不相同。搜索策略可以有两种基本类型:一种方法是从一个起始的URL集出发,顺着这些URL中的超链接,以深度优先或宽度优先,

以及启发式循环地发现新的信息。这些起始的URL集可以是任意的URL,但更多的是流行的和包含着很多链接的站点。另一种方法是将Web空间按照域名、IP地址划分,每个搜索器负责一个子域进行遍历搜索。

2）索引器

索引器的功能是理解搜索器获取的信息,进行分类并建立索引,存放到索引数据库或目录数据库中。索引数据库可以使用通用的大型数据库,如Oracle或Sybase,也可以是自己定义的文件格式。索引项可以分为两种:客观索引项与内容索引项。其中,客观索引项与文档的语意内容无关,如作者名、URL、更新时间、编码、长度与链接流行度等。内容索引项反映的是文档内容,例如关键字、权重、短语与单字等。内容索引项可以分为单索引项与多索引项(或短语索引项)。英文单索引项是单个英文单词,而中文需要对文档进行词语切分。

用户查询过程只能对索引进行检索,而不是对原始数据进行检索。索引器在建立索引时,需要为每个关键字赋予一个等级值或权重,表示该网页的内容与关键词的符合程度。当用户输入一个或一组关键词时,搜索器将查询索引数据库,找出与关键字相关的所有网页。有时被查出的网页数量很大,搜索器按照等级值由高到低排序,将排序的结果提高给用户。因此,检索结果是否符合用户的需求,取决于索引器确定关键字及权重的策略。

3）检索器

检索器的功能是根据用户输入的搜索关键字,在索引库中快速检索出文档。根据用户输入的查询条件,对搜索结果的文档与查询的相关度进行计算和评价。有的搜索引擎是在查询之前已经计算网页的等级。根据评价意见,对输出的查询结果进行排序,将相关度或等级高的排在前面,将相关度或等级低的排在后面。很多搜索引擎都具备处理用户反馈的能力。

4）用户接口

用户接口用于输入查询要求,显示查询结果,提供用户反馈意见。一个好的用户接口采用人机交互的方法,以适应用户的思维方式。用户接口可以分为两类:简单接口与复杂接口。其中,简单接口只提供用户输入关键字的界面,而复杂用户接口可以对用户输入条件进行限制,例如,进行简单的与、或、非等逻辑运算,以及相近关系、范围等限制,以提高搜索结果的有效性。

问题7-22:即时通信协议是如何发展起来的?

目前,很多即时通信系统都是采用服务提供商自己制定的即时通信协议,例如微软制定的MSNP、AOL制定的OSCAR协议、QQ制定的专用协议。由于各个公司制定的协议互相不兼容,因此不同即时通信系统之间无法实现互联互通。1999年,IETF提出了会话初始化协议(Session Initiation Protocol,SIP)。RFC3261～RFC3266文档对SIP进行了详细的描述。

SIP是在应用层实现即时通信的控制信令协议。在SIP中,“会话”是指用户之间的数据传输。传输的数据可以是普通文本数据,可以是音频或视频数据、E-mail、聊天、游戏等数据。SIP用于创建、修改和终止会话。在传输层,SIP可以使用TCP、UDP或流控制传输协议(Stream Control Trasmission Protocol,SCTP)。

问题7-23:即时通信有哪几种工作模型?

即时通信工作模型可以分为两种类型:在线的对等通信方式、离线的中转通信方式。

图 7-8 给出了典型即时通信系统 QQ 的通信过程。从图中可以看出,QQ 属于集中式的 P2P 结构。QQ 用户需要通过在线、手机、电子邮件等申请办法,在 QQ 服务器上注册并获得自己的用户名与密码。当用户需要加入 QQ 网络时,首先在自己的计算机上运行 QQ 客户端软件,然后输入自己的用户名与密码。服务器在验证客户的合法身份之后,用户就可以加入 QQ 网络中。在登录成功后,QQ 用户可以通过服务器下载自己的好友列表、在线信息,以及一些好友发送给他的离线信息。

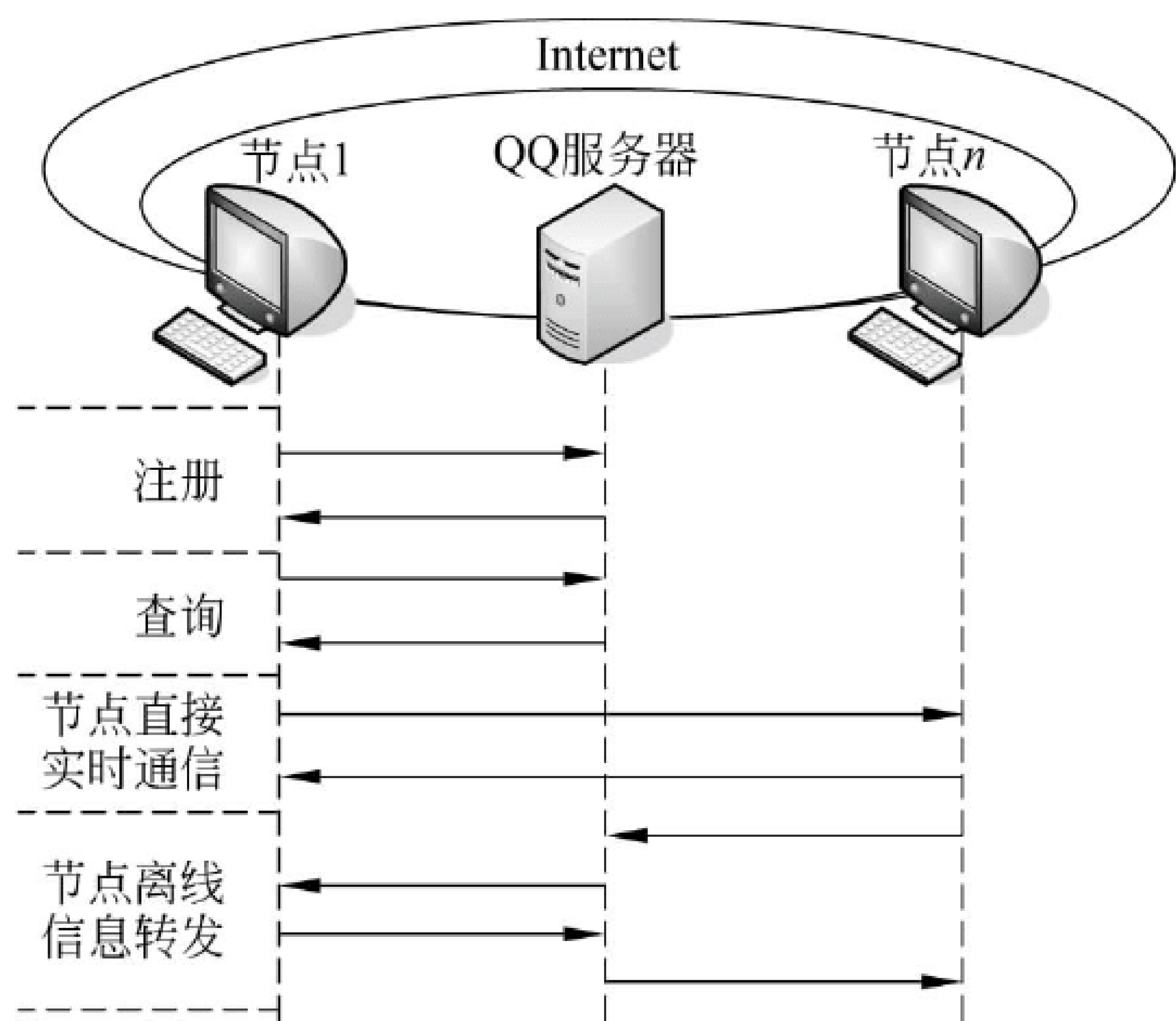


图 7-8　典型即时通信系统 QQ 的通信过程

QQ 用户之间的通信有两种方式:在线的实时的对等通信和离线的中转通信。在线信息包括希望通信的好友的 IP 地址等信息。在得到这些信息之后,用户之间可以进行点-点、直接、实时、对等的通信。离线方式可以通过 QQ 服务器间接转发信息。如果其他用户给该用户发送离线信息,用户登录 QQ 服务器时会收到离线转发的信息。

问题 7-24:SIP 具有什么样的特点?

目前,SIP 已经成为互联网的建议标准[W-SIP]。早期应用于网络电话的最主要的通信协议是 H.323,它是于 ITU-T1996 年制定的。1998 年,该协议第 2 版的名称为"基于分组的多媒体通信系统"。H.323 是关于互联网实时语音与视频会议的一组协议标准的统称,它定义了系统和组成、呼叫模式、呼叫信令过程、控制报文结构、多路复用、语音编码器、视频编码器等关键技术,因此 H.323 协议很复杂。与 H.323 协议相比,SIP 具有以下主要的特点。

1. 协议简洁,效率高

与 H.323 协议相比较,SIP 将 VoIP 作为一种新的互联网应用来处理,因此它只涉及网络电话的信令与服务质量问题,并且不规定系统一定要采用特定的语音编码器和实时传输协议(Real-time Trasport Protocol,RTP)、实时传输控制协议(RTP Control Protocol,RTCP)。因此,SIP 的结构与内容简洁、效率高。当然,在实际的网络电话系统的设计中,设计者可能需要选择 RTP 或 RTCP 作为配合的协议,但是 SIP 对于这点没有做限定,这就给应用系统的设计者很大的自由度和选择空间。

2. 客户/服务器工作模式

SIP 采用了客户/服务器工作模式，它定义了两种构件与两种状态的代理。这两种构件分别为：用户代理(User Agent，UA)与网络服务器(Network Server)。

1）用户代理

用户代理包括两个程序：用户代理客户(User Agent Client，UAC)与用户代理服务器(User Agent Server，UAS)。用户代理客户发起呼叫，而用户代理客户则接受呼叫。用户代理客户的表现形式有多种，有些是运行在计算机上的软件，有些是嵌入到移动设备(例如笔记本、PDA 或移动电话)的应用软件。

2）网络服务器

SIP 定义了三类网络服务器：代理服务器、注册服务器与重定向服务器。

代理服务器接受用户代理客户发出的呼叫请求，并将它转发给被叫用户或下一跳的代理服务器，然后由下一跳的代理服务器将呼叫请求转发给用户代理服务器，因此代理服务器也称为 SIP 路由器。

注册服务器接收和处理用户代理请求，完成用户地址注册过程。注册服务器保存用户地址与当前所在位置的映射关系。

重定向服务器不接受用户呼叫请求，只处理 SIP 呼叫路由。当它接收到代理服务器呼叫路由请求时，它通过响应报文告诉下一跳代理服务器的地址。代理服务器根据该地址重新向下一跳的代理服务器发送呼叫请求报文。

3）代理服务器的两种状态

针对代理服务器，SIP 定义了两种状态：有状态代理与无状态代理。其中，有状态代理服务器保存接收到的用户代理接入请求、回送的响应，以及转发的请求信息。无状态代理服务器在转发请求信息之后不保留状态信息。

两者相比之下，有状态代理服务器可以并行地建立和维护多个会话连接；而无状态代理服务器由于不保存用户代理请求与转发信息，因此系统响应速度会比较高。SIP 主干部分的代理服务器多采用无状态代理方式。

3. 地址灵活

SIP 使用的地址可以是电话号码，也可以是电子邮件地址或 IPv4 地址。SIP 要求的地址格式可以为：

(1) 电话号码：sip:wugongyi@8622-23508917。

(2) IPv4 地址：sip：wugongyi@202.1.2.180。

(3) 电子邮件地址：sip:wugongyi@nankai.edu.cn。

SIP 为了保证用户在移动过程中都能够进行通信，SIP 系统设置了注册服务器。用户在移动过程中向注册服务器发送信息，注册服务器不断地更新用户 SIP 地址与新位置信息的映射关系。

问题 7-25：为什么要研究动态主机配置协议 DHCP?

理解动态主机配置协议 DHCP 研究的背景，需要注意以下几个问题。

(1) 对于 TCP/IP 网络来说，要将一台主机接入 Internet 中必须配置以下参数。

① 本地网络的默认路由器地址。

② 主机应该使用的网络掩码。

③ 为主机提供特定服务的服务器地址,例如DNS、E-mail服务器。

④ 本地网络的最大传输单元MTU长度值。

⑤ IP分组的生存时间TTL值。

每台接入的主机配置的参数有十多个,只有IP地址各不相同,而其他参数应该是相同的。主机参数配置不但需要在组网时进行,在有主机加入和退出时也需要进行。作为一个网络管理员,在管理十几台主机的局域网时,主机配置任务通过手工的方法完成是可行的。但是,如果他管理的局域网接入主机的数量达到几百台时,并且经常有主机接入和移动,那么通过手工方法完成将是效率很低和容易出错的。同时,对于远程主机、移动设备、无盘工作站和地址共享配置,手工方法是不可能完成的。因此,对于大规模的网络以及远程主机、移动设备、无盘工作站和地址共享配置,用手工进行主机配置已经不可能实现,使主机参数配置过程自动化,研究动态主机配置协议就成为一个重要的问题。

(2) 动态主机配置协议可以为主机自动分配IP地址及其他一些重要的参数。动态主机配置协议不但运行效率高,减轻网络管理员的工作负担,更重要的是能够支持远程主机、移动设备、无盘工作站的地址共享与配置。

(3) 在讨论动态主机配置协议时,人们自然会想到这是网络层的任务,动态主机配置协议应该作为网络层协议之一。早期存在有硬件配置低的计算机,它的引导程序存放在PROM中,不保存网络配置参数。每一次开机接入网络时,无盘工作站引导程序需要从服务器下载配置参数。1984年研究的反向地址解析协议RARP是第一个试图解决无盘工作站引导问题的协议,它确实是放在网络层。因为RARP是通过将主机的网络层IP地址与数据链路层硬件地址绑定的方法,连接在局域网中的无盘工作站用MAC层广播的方式向RARP服务器发送请求,RARP服务器返回的应答中包含着对应该无盘工作站的IP地址。RARP机制的好处是:协议简单,易于实现。它的缺点是:一种RARP软件不可能适用不同类型的局域网,适用于Ethernet的RARP的软件就不适用于Token Bus或Token Ring,因为它们的地址与帧格式都不相同;每个局域网都需要部署一个RARP服务器;每台RARP服务器都需要人工完成地址配置表。同时,RARP服务器不能提供其他无盘工作站启动所需要的参数。

(4) 代替RARP的引导协议(Bootstrap Protocol,BOOTP),以及在BOOTP基础上发展起来的将动态主机配置协议(Dynamic Host Configuration Protocol,DHCP),它们都被放在应用层。其理由主要有两点:一是将动态主机配置协议放在应用层,可以使协议操作不依赖于低层的硬件;二是能在网络之间传送主机配置文件,这点是网络层无法实现的。

问题7-26:DHCP经历了怎样的发展过程?

首先用于TCP/IP网络的主机配置协议是引导协议BOOTP。BOOTP克服了RARP的许多缺点,支持主机配置。

1985年9月,RFC951对BOOTP进行了标准化。但是,BOOTP是一种静态配置协议。20世纪90年代,对动态IP地址分配的需求变得十分突出,这种需求导致了动态主机配置协议DHCP的出现。

BOOTP替代了RARP,而DHCP是建立在BOOTP的基础上。RFC1533将厂商BOOTP的基础上扩展的内容与DHCP融合成一个标准,形成了TCP/IP主机配置协议的标准。1997年发布的RFC2131、RFC2132是DHCP的协议草案。近年来又出现新的关于

DHCP 的协议文档，如 RFC3396、RFC3442 等。DHCP 提供一种“即插即用联网”机制，它允许一台主机接入网络之后就可以自动获取一个 IP 地址与相关的参数。同时，DHCP 可以给各种服务器分配一个永久的 IP 地址。

问题 7-27：DHCP 服务器的主要功能是什么？

DHCP 最重要的创新点是在动态 IP 地址分配与地址租用的概念上。DHCP 是基于客户/服务器工作模式。DHCP 服务器是一个为客户计算机提供动态主机配置服务的网络设备。DHCP 服务器的功能主要如下。

1. 地址储存与管理

DHCP 服务器储存 IP 地址，记录哪些 IP 地址已经被使用，哪些 IP 地址仍然可用。

2. 配置参数的储存和管理

DHCP 服务器储存和维护其他的主机配置参数。

3. 租用管理

DHCP 服务器用租用方式将 IP 地址动态地分配给主机，并管理 IP 地址的租用期。DHCP 服务器维护批准租用给主机的 IP 地址信息，以及租用期长度。RFC1533 规定租用期用 4 个字节的二进制数来表示，单位为秒。

4. 响应客户主机请求

DHCP 服务器响应主机发送的请求分配地址、传送配置参数，以及租用的批准、更新与终止等各种类型的请求。

5. 服务管理

DHCP 服务器允许管理员查看、改变和分析有关地址、租用、参数等，以及与 DHCP 服务器运行相关的信息。

问题 7-28：DHCP 客户的主要功能是什么？

DHCP 客户主机的功能主要如下。

1. 发起配置

DHCP 客户主机可以随时向 DHCP 服务器发起获取 IP 地址与配置参数的协商过程。

2. 配置参数管理

DHCP 客户主机可以从 DHCP 服务器获取全部或部分配置参数，并维护配置参数。

3. 租用管理

DHCP 客户主机可以更新租用期，在无法更新时进行重绑定，在不需要时提前终止租用。

4. 报文重传

DHCP 采用不可靠的 UDP，DHCP 客户主机要负责检测 UDP 报文是否丢失，以及丢失之后的重传。

问题 7-29：网络管理功能应该包括哪些内容？

按照 ISO 有关文档的规定，网络管理被分为 5 个部分：配置管理、性能管理、记账管理、故障管理和安全管理。

1. 配置管理

配置管理功能是监控网络中各个设备的配置信息，包括网络拓扑结构、各个设备与链路

的互连情况、每台设备的硬件和软件配置数据,以及网络资源的分配。

2. 性能管理

性能管理功能是测量和监控网络运行的状态,监视、收集和统计网络运行性能的数据,发现某个参数的当前值超过管理人员预先设定的阈值,及时通知管理人员。通过对一段时间内收集的数据的统计分析,帮助管理人员了解路由器的CPU与内存利用率、各个接口带宽利用率与输入输出I/O吞吐率、响应时间等参数。

3. 记账管理

记账管理是测量和收集各种网络资源的使用情况,统计、分析节点发送和接收的流量与使用的时间,为按流量或时间的计费提供依据。

4. 故障管理

故障是指有可能导致网络出现部分或全部中断或瘫痪,必须予以修复的错误。故障管理功能包括故障检测、差错跟踪,故障检测日志、产生报告与隔离定位。

5. 安全管理

为了保障网络正常工作,必须采取多项安全控制措施。安全管理功能是通过设定若干规则,防止网络遭受有意或无意的破坏,同时限制对敏感资源的未经授权的访问。安全管理包括:建立访问权限和访问控制;建立安全审计,对系统中各种重要操作与违规操作进行记录;当出现安全事件时发出警告和产生安全报告。

问题7-30:网络管理系统是由哪几个部分组成的?

理解网络管理系统的结构需要注意以下几点。

网络管理的目的是:使网络资源能得到有效的利用,网络出现故障时能及时报告和处理,以保证网络能够正常、高效地运行。网络管理系统通常由5个部分组成:管理进程、被管对象、代理进程、管理信息库和网络管理协议。图7-9给出了网络管理系统结构示意图。

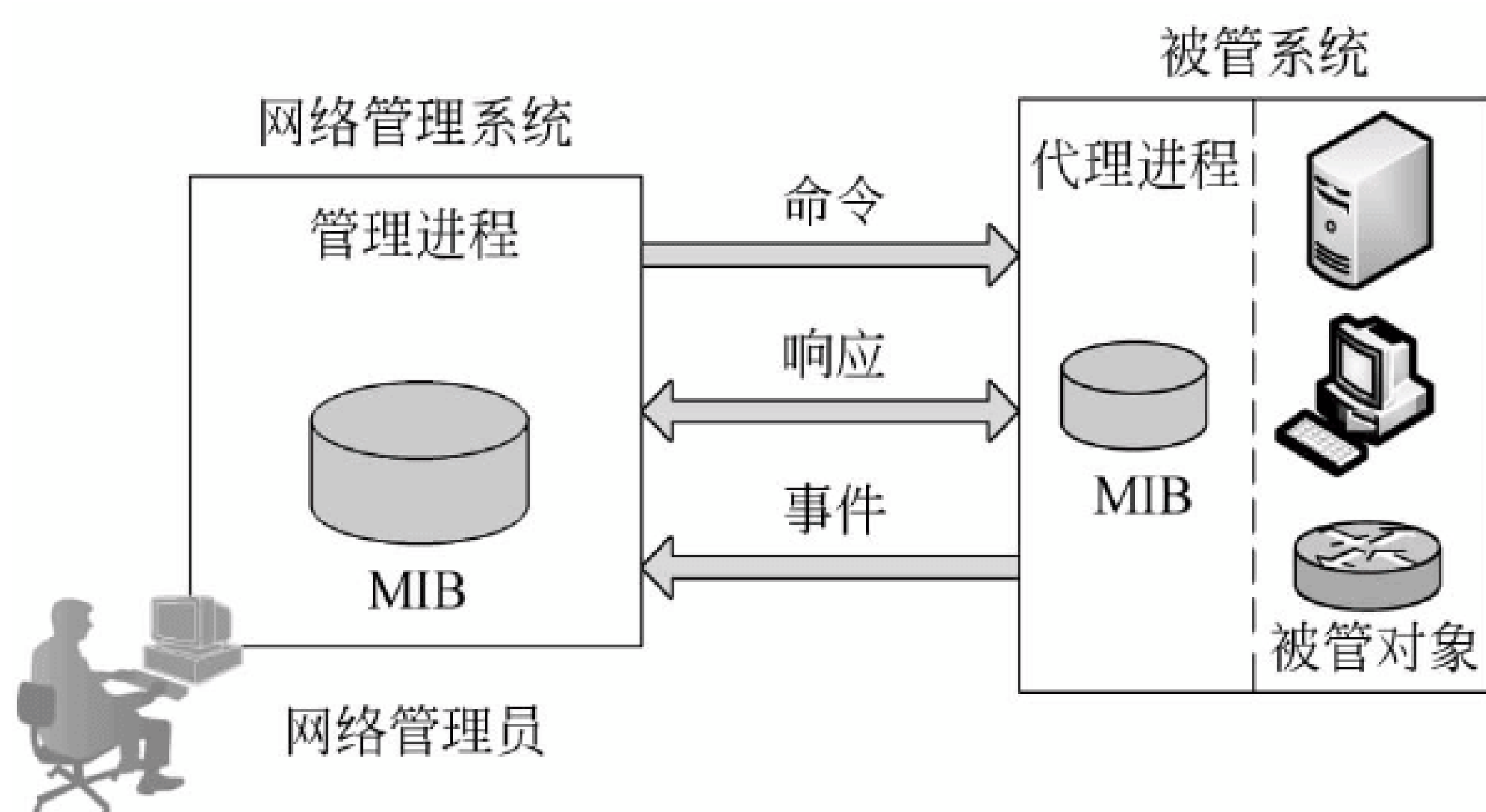


图7-9 网络管理系统结构示意图

网络管理系统的组成部分如下。

1. 管理进程

管理进程是网络管理的主动实体,它提供网络管理员与被管对象之间的界面,完成网络管理员指定的各项管理任务,读取或改变被管对象的网络管理信息。

2. 被管对象

被管对象指网络上的软硬件设备,例如交换机、路由器、主机与服务器等。

3. 代理进程

代理进程执行管理进程(例如系统配置、数据查询)的命令,向管理进程报告本地出现的异常情况。

4. 网络管理协议

网络管理协议规定了管理进程与代理进程之间交互的网络管理信息的格式、意义与过程。目前,流行的网络管理协议主要包括:TCP/IP 体系的简单网络管理协议(Simple Network Management Protocol,SNMP)与 OSI 参考模型的公共管理信息协议(Common Management Information Protocol,CMIP)。

5. 管理信息库

被管对象的信息都存放在管理信息库(Management Information Base,MIB)中。管理信息库是一个概念上的数据库。本地管理信息库只需包含与本地设备相关的信息。代理进程可以读取和修改本地 MIB 中的各种变量值。每个代理进程管理自己的本地 MIB,并与管理进程交换网络状态信息。多个本地 MIB 共同构成整个网络的 MIB。

问题 7-31:如何理解 SNMP 名称中"简单"的含义?

实际上网络管理是一个很困难的问题,它受到网络拓扑、网络规模、网络设备类型、网络状态的动态变化等因素的影响,因此描述网络管理的模型和协议也一定很复杂。网络中任何硬件与软件的增删都要影响到网络管理对象的变化,那么网络管理系统设计一定要考虑到如何将这种对象"添加"的影响减到最小。

从 SNMP 名称上可以看出,设计者希望用"简单"的系统结构和协议解决复杂的网络管理问题。"简单"应该理解为协议设计者的设计目标和技术路线。从 SNMP 的基本内容上看,SNMP 的交互过程简单,只规定 5 种消息对网络进行管理。为了简化和降低通信代价,它在传输层采用简单的 UDP。

问题 7-32:通过对 FTP 的解析,如何理解应用层协议包括哪些内容?

通过对 FTP 执行过程的解析,可以对应用层、网络服务功能与应用层协议的关系有更具体和深入的理解,同时也可以知道应用层协议应该包括哪些内容。FTP 工作模型对于需要保证数据可靠传输的应用有很好的借鉴意义。

1. 应用层协议的基本概念

网络应用与应用层协议是两个重要的概念。E-mail、FTP、TELNET、Web、IM、IPTV、VoIP,以及基于网络的金融应用系统、电子政务、电子商务、远程医疗、远程数据存储都是不同类型的网络应用。应用层协议规定了应用程序进程之间通信所遵循的通信规则,包括:如何构造进程通信的报文,报文应该包括哪些字段,每个字段的意义与交互的过程等问题。

以 Web 服务为例,Web 网络应用程序包括 Web 服务器程序、Web 浏览器程序。Web 应用层协议(HTTP)定义了 Web 浏览器与 Web 服务器之间传输的报文格式、会话过程与交互顺序。

对于电子邮件应用系统来说,电子邮件应用程序包括邮件服务器程序与邮件客户端程序。电子邮件应用层协议(SMTP)定义了服务器与服务器之间、服务器与邮件客户端程序之间传送报文的格式、会话过程与交互顺序。

对于文件传输服务来说,FTP 规定文件传输服务系统的结构、工作模式、进程会话连接

的服务器端的熟知端口号,以及会话过程的顺序与应答关系。

2. 应用层协议的基本内容

应用层协议定义了运行在不同端系统上应用程序进程交换的报文格式与交互过程,它主要包括以下内容。

(1) 交换报文的类型,如请求报文与应答报文。

(2) 各种报文格式与包含的字段类型。

(3) 对每个字段意义的描述。

(4) 进程在什么时间、如何发送报文,以及如何做出响应。

问题7-33:如何理解应用层协议的分类?

根据应用层协议在Internet中的作用和提供的服务功能,作者认为应用层协议可以分为三种基本类型:基础设施类、网络应用类与网络管理类。图7-10给出了主要应用层协议分类的示意图。

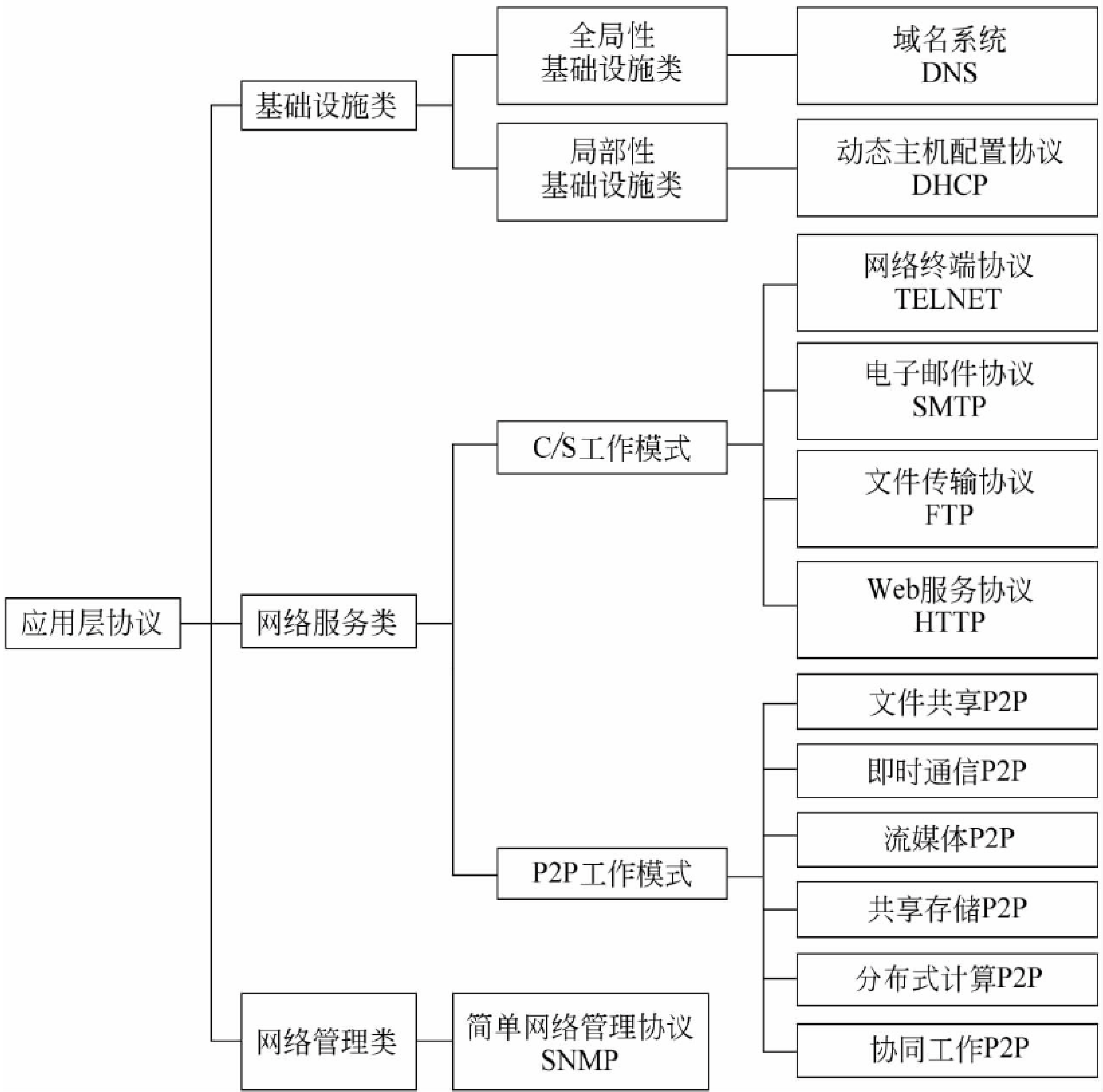


图7-10 应用层协议分类

1. 基础设施类

属于基础设施类的应用层协议主要有以下两种。

(1) 支持Internet运行的全局基础设施类应用层协议——域名服务协议(DNS)。

(2) 支持各个网络系统运行的局部基础设施类应用层协议——动态主机配置协议(DHCP)。

2. 网络应用类

网络应用类的协议可以分为两类：基于C/S工作模式的应用层协议与基于P2P工作模式的应用层协议。

1）基于C/S工作模式的应用层协议

基于C/S工作模式的应用层协议主要包括网络终端协议TELNET、电子邮件服务的SMTP、文件传输服务的FTP、Web服务的HTTP等。

2）基于P2P工作模式的应用层协议

目前，很多P2P都属于专用应用层协议。P2P基本上可以分为：文件共享P2P、即时通信P2P、流媒体P2P、共享存储P2P、协同工作P2P。

3. 网络管理类

网络管理类的协议主要有简单网络管理协议SNMP。

问题7-34：如何理解网络应用与各层协议之间的关系？

在完成一种典型的Internet应用层协议剖析之后，有必要回顾和总结网络应用层协议与低层协议之间的关系。图7-11从接收到一个Ethernet帧开始，主机通过各层协议的分解与协同，完成一次主机之间的进程通信过程的描述，试图以一个简化的过程来对一个复杂的协议之间的关系做一个总结。虽然这不是TCP/IP全部的工作过程，但是它能很好地体现出计算机网络分层结构的设计思想。

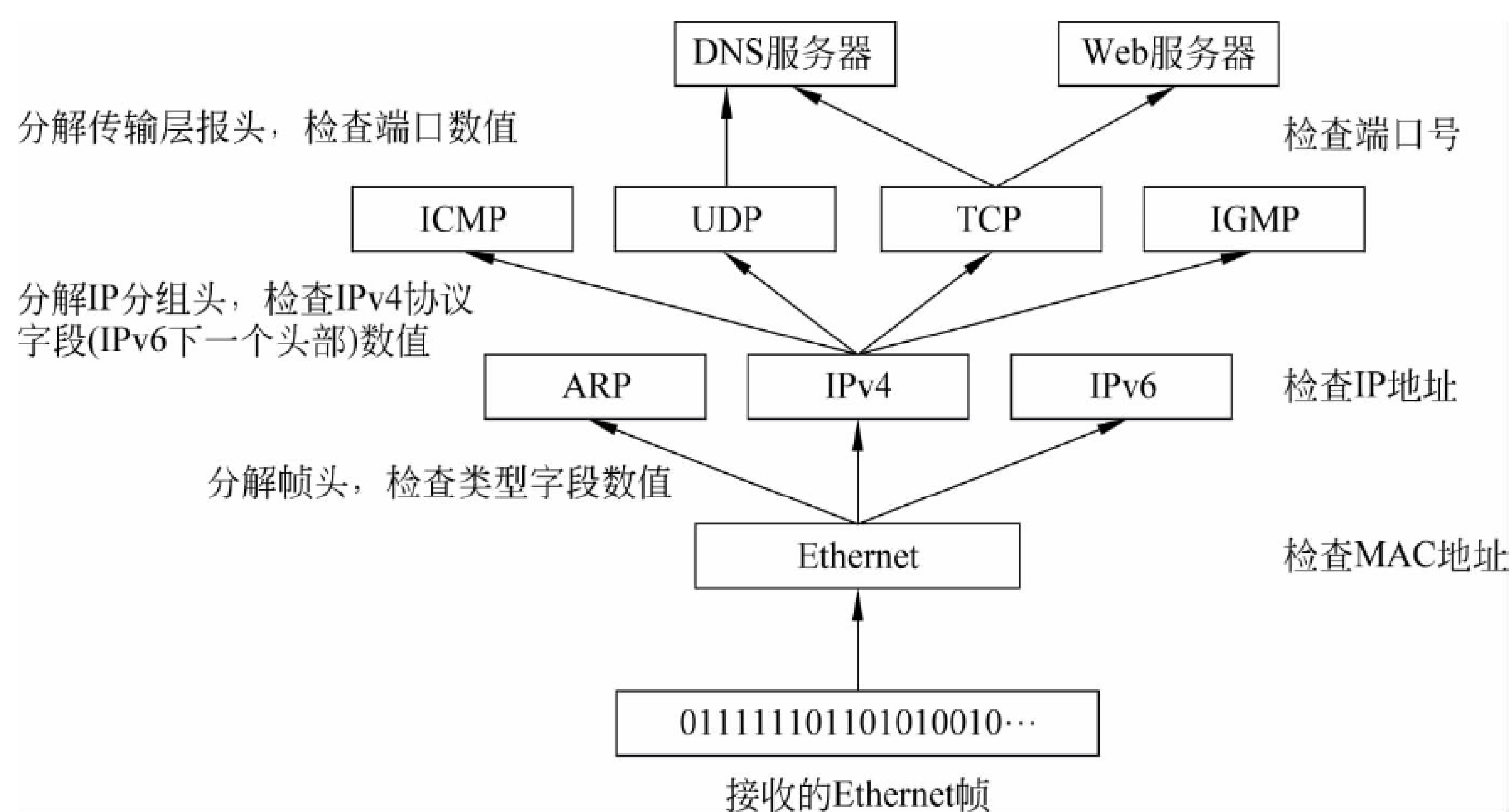


图7-11　网络应用与各层之间的关系示意图

当一台主机接收到一个Ethernet帧时，这个帧的帧头包含着一个长度为48b的目的地址（MAC地址），以及一个16b的协议类型字段。类型字段表示高层网络层所使用的协议类型。例如，类型字段值等于0x0800，表示网络层使用IPv4协议；类型字段值等于0x8106，表示地址解析ARP；类型字段值等于0x86DD，表示网络层使用IPv6协议。

如果帧的类型字段值是“0x0800”或“0x86DD”，表示接收的Ethernet的数据字段中封装了一个IPv4分组或IPv6分组。那么，下一步的工作就是检查目的地址，32b的IPv4或128b的IPv6地址，确定目的地址是否匹配。如果匹配，说明这是发送给本主机的IP分组；IP协议软件将根据IP分组头的8b的IPv4的协议字段，确定接下来调用哪种传输层协议

来继续处理。协议字段是指使用IP协议的高层协议类型。IPv4的协议字段长度为8位，协议字段数值为1，表示高层协议是ICMP；数值为6，表示高层协议是TCP；数值为17，表示高层协议是UDP；数值为4，表示高层协议是IPv6协议。看起来很奇怪，这违反了分层与封装的原则，但是它是作为隧道技术的基础。当然，这里简化了数据正确性的检查步骤。

如果高层是TCP或UDP，那么就需要检查端口号。端口号是0～65 535范围内的数值。对于客户/服务器工作模式，一台服务器首先要"绑定"一个端口号，然后一个或多个客户机可以使用一种协议，在自己的端口号与这个服务器的端口号上建立传输连接。客户机的进程与服务器的进程通过这个传输连接来交换数据，享受服务器提供的服务功能。

问题7-35：如何理解应用层网络应用软件设计与开发方法？

总结第7章的学习，我们通过对应用层、应用层协议、网络服务功能实现方法的分析，可以形成对网络应用软件设计与开发方法的基本看法。网络应用软件设计、开发的基本方法大致可以分为以下几步。

(1) 根据网络应用的功能要求，设计相应应用层协议的工作模型。协议工作模型描述了为了实现服务功能的客户、服务器双方的基本结构与模块组成，以及对服务质量的要求。根据协议工作模型选择和确定各层的协议类型。

(2) 根据协议工作模型，确定应用层实体的各个模块之间信息交互的时序与内容，设计协议数据单元的结构，并以此作为系统实现与软件编程的依据。

(3) 软件开发人员在理解协议模型、读懂协议规定的基础上完成编程任务。

第三部分　习题参考答案

1. 应用层协议与对应的低层协议名称为：

	Web服务	网络管理服务	虚拟终端服务	电子邮件服务	动态主机地址分配服务	域名服务	文件传输服务
应用层	HTTP	SNMP	TELNET	SMTP	DHCP	DNS	FTP
传输层	TCP	UDP	TCP	TCP	TCP	UDP	TCP
网络层	IP	IP	IP	IP	IP	IP	IP

2. (1) 050122450066
   (2) 212.8.2.28
   (3) 255.255.255.0
   (4) 212.8.20.2
   (5) 2011-06-20 09:06:05
   (6) 2011-06-28 09:06:05
3. (1) DNS服务器的IP地址是131.1.64.16。
   (2) 图中删除的①的信息是ACK。
   (3) 主机202.1.2.197是FTP服务器；使用的熟知端口号是21。
   (4) 访问FTP服务器的主机使用的临时端口号是59088。

4.（1）该主机的 IP 地址是 202.1.64.166。

（2）该主机正在浏览的网站是 www.nk.edu.cn。

（3）该主机设置的 DNS 服务器的 IP 地址是 211.80.20.200。

（4）该主机用 HTTP 通信时使用的源端口号是 1535。

（5）TCP 连接三次握手过程完成的报文号是 No.7。

5. 服务器对象标识符为 1.3.6.1.4.150.50。

6.（1）Web 服务器的 IP 地址是 64.170.98.32；主机 A 的默认网关的 MAC 地址是 00-21-27-21-51-ee。

（2）在构造 ARP 请求包的帧中，目的 MAC 地址用 ff-ff-ff-ff-ff-ff，以广播的方式发送该帧。

（3）从发出请求到收到全部内容，需要经过 6 个 RTT。

（4）IP 分组经过路由器 R 转发时，需要修改 IP 分组头中的生存时间（TTL）与校验和字段。

第 8 章 网络安全

第一部分 学习目的、要求与知识点结构

1. 学习目的

本章将系统地学习网络安全的基本概念、密码体制的基本概念、网络安全协议、防火墙、入侵检测技术、网络业务持续性规划技术、网络防病毒技术，以及网络管理技术。通过本章的学习，读者将初步建立网络安全的基本概念，掌握密码体制、防火墙、入侵检测、网络文件的备份与恢复、网络防病毒与网络管理的技术要点。

2. 学习要求

(1) 了解：网络安全的重要性。

(2) 了解：当前网络安全形势的变化。

(3) 理解：密码体制的基本概念及应用。

(4) 掌握：网络安全协议的概念及应用。

(5) 掌握：防火墙的基本概念。

(6) 掌握：网络入侵检测与防攻击的基本概念与方法。

(7) 掌握：网络业务持续性规划技术的基本概念与方法。

(8) 理解：网络病毒防治的基本概念与方法。

3. 本章知识点的组织与结构

本章知识点的组织与结构如图 8-1 所示。

第二部分 教学内容问答

问题 8-1：如何认识网络安全技术的特点？

网络安全技术的特点体现在以下三个方面。

1. 网络安全是网络技术研究中一个永恒的主题

计算机网络与 Internet 是高悬在全人类头上的一把双刃剑。一方面，计算机网络与 Internet 的应用对于各国的政治、经济、科学、文化、教育与产业的发展起到了重要的推动作用。另一方面，人们也对它的负面影响也忧心忡忡。网络安全的研究一直是伴随着网络技术与应用的发展而进步。

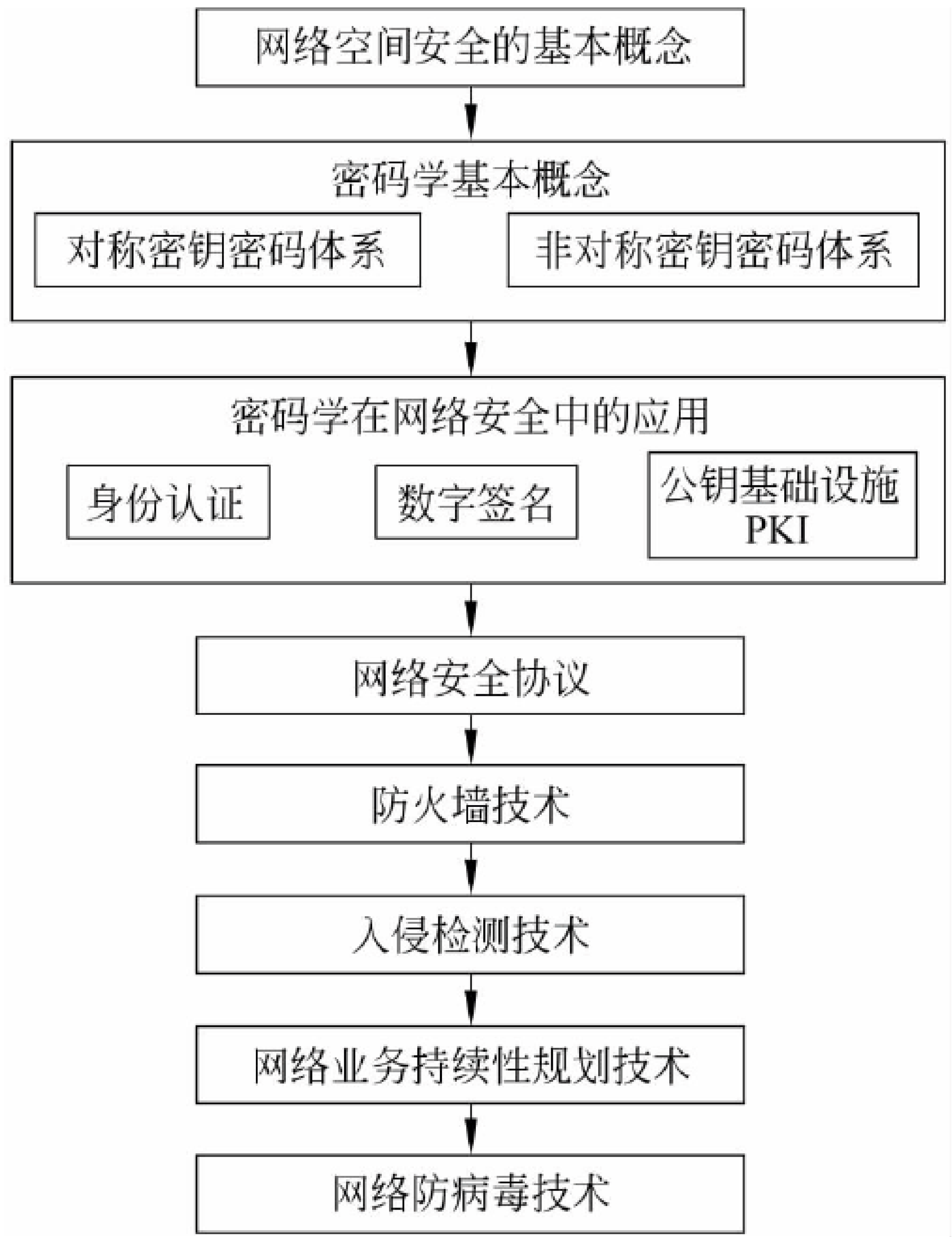


图 8-1　知识点的组织与结构

2．网络安全是一个系统的社会工程

现实社会有什么，网络虚拟社会就会有什么。这是为什么？答案其实很简单，是生活在现实世界的人创造了网络神话，也是现实世界的人在使用网络。人们会将现实世界的几乎所有的东西都“克隆”到网络的虚拟社会之中。因此，现实世界中真善美的东西，网络的虚拟社会都会有。同样，现实社会中丑陋的东西，网络的虚拟社会一般也会有，只是迟早的问题，只是表现形式不一样。图 8-2 形象地描述了这个规律。如果透过复杂的技术术语和面对的计算机用户界面，计算机网络虚拟空间和现实社会的实际空间应该是“对应”的。在现实社会中，在人与人的交往中形成了复杂的社会与经济关系。在网络社会里，这些关系以数字化的方式延续着。

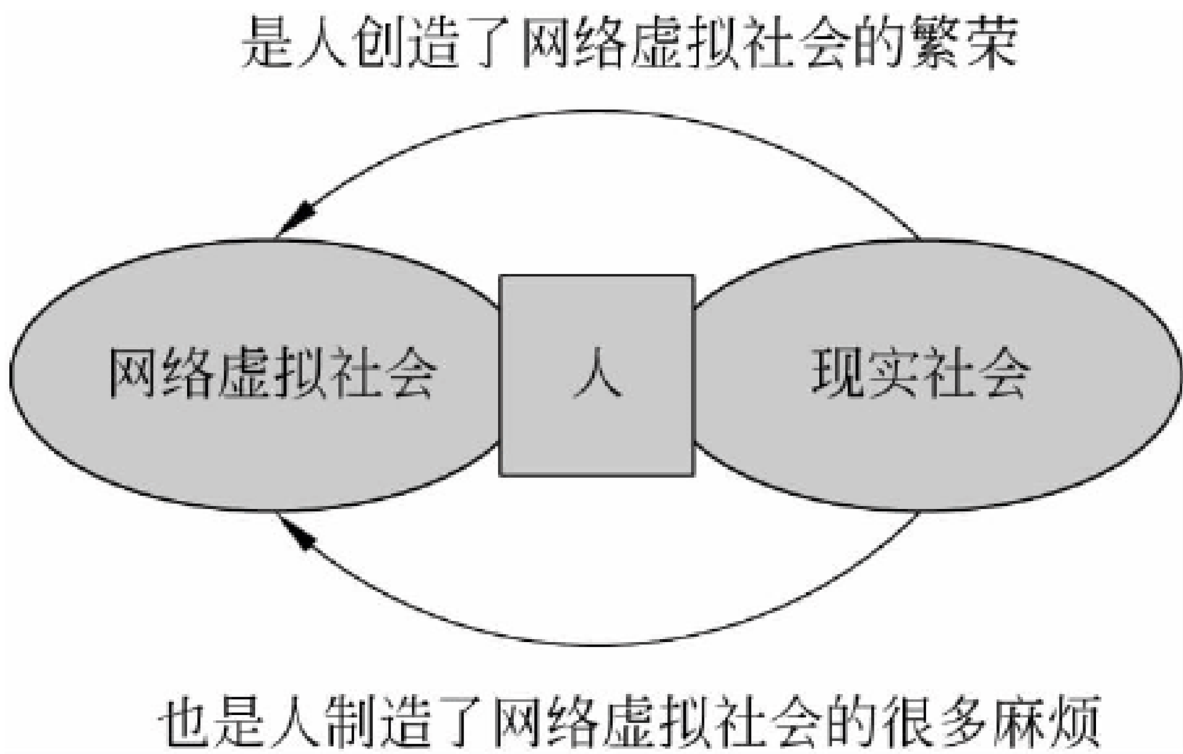


图 8-2　网络虚拟社会与现实社会的关系

网络安全是现实社会安全的反映。网络安全问题实际上是个社会问题，光靠技术来解决这些问题是不可能的。网络安全是一个系统的社会工程，它必然要涉及技术、政策、道德与法律法规。

3．网络安全危及国家安全

网络安全问题已成为信息化社会的一个焦点问题，并且已经成为网络安全危及国家安全与稳定的大问题。每个国家只能立足于本国，研究网络安全技术，培养专门人才，发展网络安全产业，才能构筑本国的网络与信息安全防范体系。哪个国家不重视网络与信息安全，他们必将在未来的国际竞争中处于被动和危险的境地。

问题 8-2：网络安全与信息系统安全是什么关系?

所有的信息系统都是建立在计算机网络与 Internet 之上。正是由于这个原因,可以说信息系统的安全都是建立在 Internet 安全的基础之上。图 8-3 给出了物联网安全与计算机、计算机网络安全关系的示意图。



图 8-3　信息系统安全与计算机网络安全的关系

信息系统都是建立在 Internet 之上。Internet 包括端系统与网络核心交换两个部分。端系统包括计算机硬件、操作系统、数据库系统等,而运行信息系统的大型服务器或服务器集群,及用户的个人计算机都是以固定或移动方式接入 Internet 中,它们是保证信息系统正常运行的基础。任何一种信息系统功能和服务的实现都需要通过网络核心交换在不同的计算机系统之间进行数据交互。病毒、木马、蠕虫、脚本攻击代码等恶意代码可以利用 E-Mail、FTP 与 Web 系统进行传播,网络攻击、网络诱骗、信息窃取可以在 Internet 环境中进行。那么,它们同样会对信息系统构成威胁。

如果 Internet 核心交换部分不安全,那么信息系统的安全就无从谈起。因此,保证网络核心交换部分的安全,以及保证计算机系统的安全,它们是保障信息系统安全的基础。

问题 8-3：网络安全与网络应用技术发展是什么关系?

按照正常人的思维方式,一位技术人员在研究和开发一种基于网络的新应用技术与系统时,只会想到这种应用可以给人们的生活和工作带来什么样的好处和乐趣,一般不会去刻意地想到黑客或居心不良的人会利用这种技术做什么坏事。而黑客恰恰是一类逆向思维和不按正常规律办事的人,他们不遵守正常人所遵循的道德规范。“Everything over IP,IP over everything”说明了计算机网络技术的成功,但它所带来的问题也是网络技术人员所始料未及的。

病毒、木马、蠕虫、脚本攻击代码的恶意代码利用电子邮件、FTP 与 Web 技术的传播,网络攻击、网络诱骗、网络欺诈,都已经是司空见惯的事了。但是这些都是在计算机网络环境中进行的。P2P 是一种十分有价值的网络应用模式,但是 P2P 除了可能方便 MP3 的盗版之外,也给恶意代码的传播提供了一种新的途径。VoIP 是计算机网络与电话服务结合的一个十分成功的应用,但是由于很多 VoIP 交换系统是在 Windows 操作系统上开发的,因此 Windows 操作系统的所有漏洞,都为黑客攻击 VoIP 交换系统提供了机会。手机病毒与无线射频标识 RFID 芯片可能感染病毒的研究结果公布,表明移动设备将成为黑客和恶意软件编写者下一个主攻的目标。

网络技术不是在真空之中,计算机网络是要提供给全世界的人使用的,网络技术人员在研究和开发一种新的基于网络应用技术与系统时,必须面对这样一个复杂的局面和现实,成功的网络应用技术与成功的应用系统的标志是功能与安全性的统一。网络安全问题不简单是部分从事网络安全技术工程师的事,也是每位信息技术领域的工程师与管理人员必须面对的问题。

问题 8-4：网络安全技术研究包括哪些基本的内容?

总结近年来网络安全研究的内容、方法与技术的发展,可以将网络安全研究的内容归纳

为如图 8-4 所示几个方面的问题。

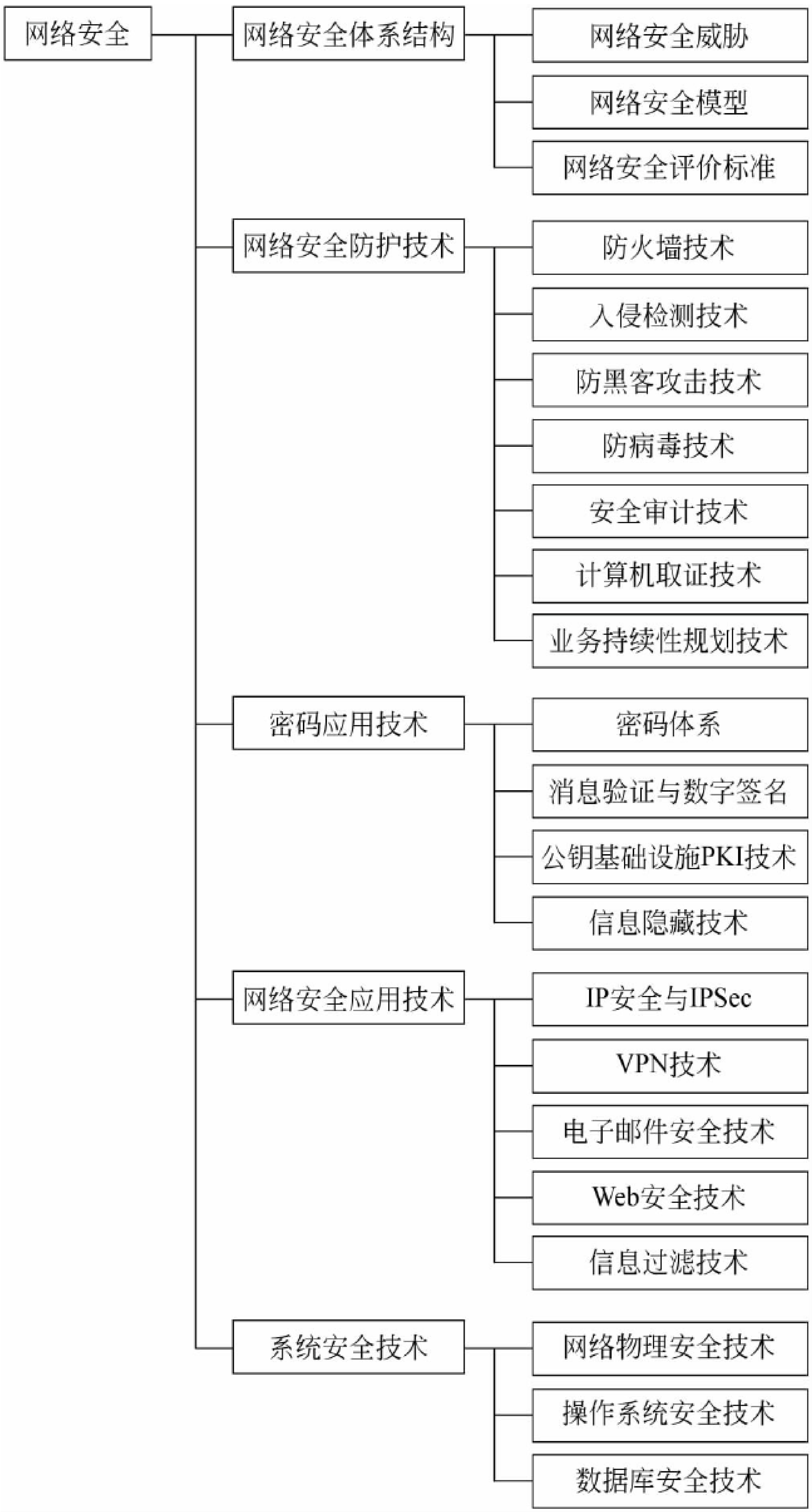


图 8-4　网络安全技术研究内容的分类

网络安全技术研究涉及以下几个方面的内容。

1．网络安全体系结构的研究

网络安全体系结构的研究主要涉及网络安全威胁分析、网络安全模型与确定网络安全体系，以及对系统安全评估的标准和方法的研究。根据对网络安全威胁的分析，确定需要保护的网络资源，对资源攻击者、攻击目的与手段、造成的后果进行分析；提出网络安全模型，并根据层次型的网络安全模型，提出网络安全解决方案。网络安全体系结构的研究的另一个重要内容是系统安全评估的标准和方法，这是评价一个实际网络应用系统安全状况的标准，是提出网络安全措施的依据。

2. 网络安全防护技术

网络安全防护技术的研究涉及防火墙技术、入侵检测技术与防攻击技术、防病毒技术、安全审计与计算机取证技术,以及业务持续性技术。

3. 密码应用技术

密码应用技术的研究涉及包括对称密码体制与公钥密码体制的密码体系,以及在此基础上研究的消息认证与数字签名技术、信息隐藏技术、公钥基础设施 PKI 技术。

4. 网络安全应用

网络安全应用技术研究主要包括 IP 安全、VPN 技术、电子邮件安全、Web 安全与网络信息过滤技术。

问题 8-5:网络安全与密码学是什么关系?

人们对密码学与网络安全的关系的认识有一个过程,这个问题可以用 Bruce Schneier 在 *Secrets and Lies: Digital Security in a Networked World* 一书的前言中讲述的观点来解释。Schneier 在 1996 年出版了一本在信息安全方面非常经典的书 *Applied Cryptography*。4 年之后他写了第二本书。他在这本书的前言中说明,他写第二本书的动机之一是为了纠正第一本书中的一个错误。他说,在第一本书中"我描述了一个数学的乌托邦:密码算法能将你最深的秘密保持数千年","密码学是超凡的技术均衡器,任何一个人只要有一台便宜的计算机,就可以达到与最强大的政府同样的安全性。"他现在认为:"事实并非如此,密码学并不能做那么多的事。"密码学并非存在于真空之中。密码学是数学的一个分支,它涉及数字、公式与逻辑。数学是完美的,而现实社会却是无法用数学去准确描述的。数学是精确的和遵循逻辑规律的,而计算机和网络安全涉及的是人所知道的事,人与人之间的关系,以及人和机器之间的关系。人是有欲望的,是不稳定的,甚至是难于理解的。Schneier 在出版第一本书之后就成了美国设计和分析一些大的信息系统的顾问。但是,后来的经历告诉他,安全性的弱点与数学"毫无关系",它们存在于硬件、软件、网络和人的身上。他认识到:"安全性是一个链条,它的可靠程度取决于链中最薄弱的环节。"同时,他认为:"安全性是一个过程,而不是一个产品。"

从一位数学家的认识转变中得到的启示是:密码学是研究网络安全所必需的一个重要的工具与方法,但是网络安全研究所涉及的问题要广得多,它是计算机、电子、通信、数学、社会学、行为学与认知心理学,以及法学的交叉学科。这也告诉人们,网络安全的研究需要计算机与电子工程师、数学家、社会学家、心理学家,以及法律界等各方面的人士参加。

问题 8-6:密钥的位数越长是不是就越好?

这个问题不可能得到一个确切的答案。这里有密码自身的安全性、加密与解密需要的计算量、时间,系统实现的造价等一系列的因素。如何一个密码系统都是在这些影响的因素之中,取一个能够被用户接受的折中方案。

对于同一加密算法,密钥的位数越长,破译的困难也就越大,安全性也就越好。在给定的环境下,为了确保加密的安全性,人们一直在争论密钥长度,即密钥使用的位数。密钥位数越多,密钥空间越大,也就是密钥的可能范围也就越大,那么攻击者也就越不容易通过蛮力攻击来破译。在蛮力攻击中,破译者可以用穷举法对密钥的所有组合进行猜测,直到成功地解密。表 8-1 给出了在给定密钥长度下,用穷举法进行猜测时需要尝试的密钥个数。

表 8-1　密钥长度与密钥个数

密钥长度/b	组合个数	密钥长度/b	组合个数
40	$2^{40}=1\ 099\ 511\ 627\ 776$	112	$2^{112}=5.192\ 296\ 858\ 535\times10^{33}$
56	$2^{56}=7.205\ 759\ 403\ 793\times10^{16}$	128	$2^{128}=3.402\ 823\ 669\ 209\times10^{38}$
64	$2^{64}=1.844\ 674\ 407\ 371\times10^{19}$		

假设用穷举法破译，猜测每 10^6 个密钥用 1μs 的时间，那么猜测 2^{128} 个密钥最长时间大约是 1.1×10^{19} 年。所以，一种自然的倾向就是使用最长的可用密钥，它使得密钥很难被猜测出。但是密钥越长，进行加密和解密过程所需要的计算时间也将越长。我们的目标是要使破译密钥所需要的“花费”比该密钥所保护的信息价值还要大。许多国家对于基于密钥长度的加密产品有着特殊的进口和出口规定。

问题 8-7：如何认识公钥基础设施 PKI 的作用？

公钥基础设施（Public Key Infrastructure，PKI）是利用公钥加密和数字签名技术建立的提供安全服务的基础设施，从而保证网络环境中数据的机密性、完整性与不可抵赖性。理解 PKI 的基本概念需要注意以下几个基本问题。

（1）PKI 是一种针对电子商务、电子政务应用，利用非对称加密密码体系实现并提供安全服务的通用性网络安全基础设施。

（2）PKI 系统对用户是透明的，用户在获得加密和数字签名服务的时候，不需要知道 PKI 是如何管理证书与密钥的。

（3）PKI 建立的安全通信信任平台与密钥管理体系，能够为所有的网络应用提供加密与数字签名服务，实现 PKI 系统的关键是密钥的管理。

（4）PKI 主要任务是确定用户可信任的数字身份。这个信任关系是通过公钥证书来实现的。公钥证书就是用户身份与所持有公钥的结合，而这种结合关系是由可信任的第三方权威机构——认证中心来确认的。

现在我们能够使用网上银行、电子商务网上结算，这一切经济活动都是家里在 PKI 的基础之上，因此掌握 PKI 的基本知识对于网络知识的学习是十分重要的。

问题 8-8：数字签名到底能够起到什么样的作用？

认识这个问题需要注意以下几点。

（1）亲笔签名是用来保证文件或资料真实性的一种方法。在网络环境中，通常使用数字签名技术来模拟日常生活中的亲笔签名。数字签名将信息发送人的身份与信息传送结合起来，可以保证信息在传输过程中的完整性，并提供信息发送者的身份认证，以防止信息发送者抵赖行为的发生。目前各国已制定了相应的法律、法规，把数字签名作为执法的依据。利用非对称加密算法（例如 RSA 算法）进行数字签名是最常用的方法。

（2）数字签名需要实现以下三项主要的功能。

① 接收方可以核对发送方对报文的签名，以确定对方的身份。

② 发送方在发送报文之后无法对发送的报文及签名抵赖。

③ 接收方无法伪造发送方的签名。

（3）目前广泛应用的数字签名算法是消息摘要 MD5（Message Digest 5）算法。它是 Rivest 于 1994 年发表的一种单向散列算法，可以对任意长度的数据生成 128 位的散列值，

也叫作不可逆指纹。攻击者不能从MD5生成的散列值反向算出原始数据。RFC1321文档对MD5做出了详细的说明。需要注意的是：MD5算法实际上没有对任何数据进行加密或修改，只是生成了一个用于判断数据完整性与真实性的散列值。

因此，利用数字签名可以验证数据在传输过程中没有被篡改，同时能够确认发送者的身份，防止信息交互中抵赖现象的发生。数据加密可以防止信息在传输过程中被截获，但是如何确定发送人的身份问题，就需要使用数字签名技术来解决。

问题8-9：网络安全协议具有哪些特点？

随着计算机网络技术在政府、军事、商务领域的广泛应用，网络协议的安全性一直是各国政府与产业界研究的重点。网络安全协议借助于密码算法达到密钥分配、身份认证、信息保密与安全传输的目的。网络安全协议是网络安全研究领域一个重要的概念。看一个协议是否称得上是一个安全的网络协议，需要注意以下4个方面的特性。

1. 认证性

认证是网络系统中的进程通信实体之间身份识别、建立信任关系的过程。认证是最重要的安全性之一，也是其他安全性的基础。

2. 秘密性

秘密性的目的是保护协议交换过程中的消息不被泄露，以及不被攻击者观测到消息的格式和含义。实现秘密性的主要方法是对协议交互过程中出现的消息进行加密。

3. 完整性

完整性的目的是保护协议交换过程中的消息不被非法篡改、删除和替代。实现完整性的主要方法是对协议交互过程中出现的消息进行封装或签名。

4. 不可否认性

不可否认性的目的是提供足够的证据，使得协议通信的主体必须对自己合法的行为负责，不能也无法事后否认，以保护协议双方的合法利益不受侵害。

同时，需要注意：网络安全协议的研究与标准的制定涉及网络层、传输层与应用层等多个层次。

问题8-10：如何认识网络层安全协议IPSec的特点？

认识IPSec协议的特点需要注意以下几个基本的问题。

1. IPSec安全体系结构

通过讨论IPv4协议可以看出：IP协议本质上是不安全的，伪造一个IP分组、篡改IP分组的内容，窥探传输中的IP分组的内容都是比较容易的。接收端不能保证每一个IP分组源地址的真实性，也不能保证IP分组在传输过程中没有被篡改或泄露。IP分组的校验和对于IP分组数据完整性的验证能力很弱，攻击者完全可以在修改IP分组数据之后，很方便地重新计算校验和，然后填回到校验和字段的位置。

为了解决IP协议的安全性问题，IETF于1995年成立了一个IP安全协议工程组，着手研究并提出了一系列的关于增强网络层的安全性的协议IP Security Protocol(IPSec)，构成了一个IP协议安全体系。1998年，IETF公布了Internet网络层的系列文档(RFC2401～RFC 2411)。

2. IPSec的主要特征

理解IPSec设计思路与技术特征，需要注意以下几个问题。

(1) IPSec的安全服务是在IP层提供的，可以为任何高层协议，如TCP、UDP、ICMP、BGP提供服务。

(2) IPSec不是单一的一种协议，IPSec安全体系主要是由认证头协议、封装安全载荷协议与Internet密钥交换协议等组成。

(3) 认证头(Authentication Header，AH)协议用于增强IP协议的安全性，提供对IP分组源认证、IP分组数据传输完整性与防重放攻击的安全服务。但是，AH协议并不对IP分组数据进行加密。

(4) 封装安全载荷(Encapsulating Security Payload，ESP)协议提供对IP分组源认证、IP分组数据完整性、机密性与防重放攻击的安全服务。

(5) Internet密钥交换(Internet Key Exchange，IKE)协议用于协商AH协议与ESP协议所使用的密码算法与密钥管理体制。

(6) 安全关联(Security Association，SA)是IPSec的工作基础。安全关联是建立网络层安全连接的双方，通过IKE协议协商将采用的加密与认证算法的过程。通过安全关联协商双方进行认证时使用的认证算法、密钥以及密钥生存期。

(7) IPSec定义了两种保护IP分组的模式：传输模式与隧道模式。

(8) IPSec对于IPv4是可选的，而是IPv6基本的组成部分。

问题8-11：如何理解IPSec VPN的技术特点？

理解隧道模式ESP协议的基本工作原理，就很容易理解IPSec VPN的技术特点。隧道模式工作原理如图8-5所示。

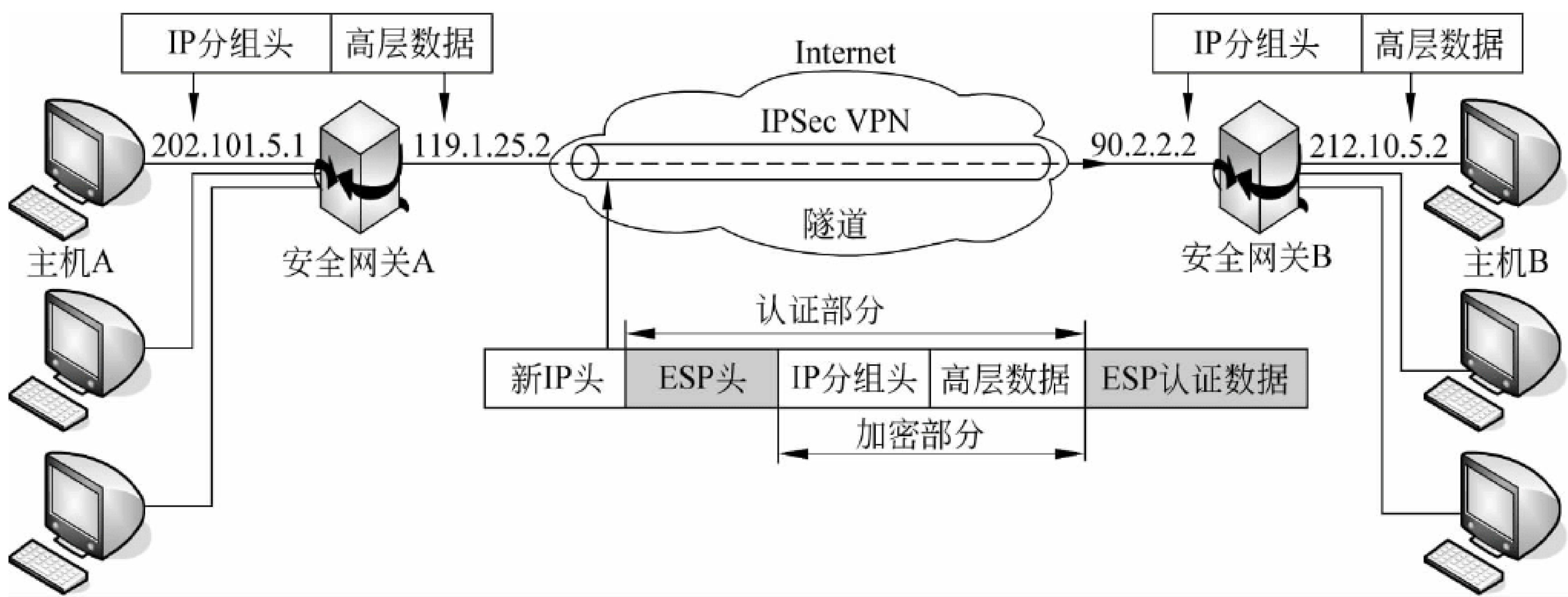


图8-5 隧道模式ESP工作原理示意图

理解隧道模式工作原理与技术特点需要注意以下几点。

(1) 隧道模式一般需要通过安全网关实现，由安全网关执行ESP协议。如果图8-5中主机A与主机B通过安全网关A与安全网关B建立网络层安全连接，ESP协议执行过程由安全网关A与安全网关B完成。这个过程对于主机A与主机B是透明的。

(2) 在隧道模式中，原始的IP分组经过安全处理之后，将被封装在新的IP分组中通过隧道传输。新的IP分组头中源与目的IP地址使用的是隧道两端安全网关的IP地址。

(3) 隧道模式一般采用ESP协议提供主机认证与IP分组数据加密服务,所采用的加密与认证算法是在安全网关建立安全关联过程中协商确定的。

(4) 对原始IP分组进行加密,可以保证分组传输的安全性;对ESP头、加密的原始IP分组进行认证,可以确认发送主机与接收主机的身份的合法性。

(5) ESP协议可以根据不同类型的应用需求,提供不同强度的加密算法,以增加攻击者破译密钥的难度,提高IP传输的安全性。

将IPSec隧道模式与构建VPN相结合,利用IPSec支持身份认证与访问控制、保证数据机密性与完整性服务,为大型网络系统在Internet环境中建立安全的IPSec VPN提供了重要的技术保证。

问题8-12:安全电子邮件协议研究的基本思路是什么?

1. 电子邮件安全的现状

电子邮件存在的垃圾邮件、诈骗邮件、炸弹邮件、病毒邮件等问题已经引起了人们高度的重视。未加密的电子邮件在网络上是很容易被截获的,如果电子邮件不是数字签名的,那么用户无法确定邮件是从哪里发送来的。

为了解决电子邮件的安全问题,主要有以下几种研究途径:端-端的电子邮件安全、传输层安全、邮件服务器安全,以及客户端的安全电子邮件技术。

2. 与电子邮件安全相关的协议与标准的研究现状

目前,已经出现不少与电子邮件安全相关的协议与标准,如:

(1) PEM(Privacy Enhancement for Internet Electronic Mail);

(2) PGP(MIME security with Pretty Good Privacy);

(3) S/MIME(Secure MIME);

(4) MOSS(MIME Object Security Services)。

3. 数字信封技术与电子邮件安全中的应用

数字信封技术用来保证数据在传输过程中的安全。在数字信封技术中需要有两个不同的加密解密过程:明文本身的加密解密与对称密钥的加密解密。首先,它使用对称加密算法对发送的明文进行加密;然后,利用非对称加密算法对对称密钥进行加密数字信封技术使用两层加密体制,在内层利用了对称加密技术,每次传送信息都可以重新生成新的密钥,保证了信息的安全性。在外层利用非对称加密技术加密对称密钥,保证密钥传递的安全性,实现了身份认证。

4. 安全电子邮件的工作模式

图8-6给出安全电子邮件工作模式的示意图。一般情况下,安全电子邮件在发送端需要加上邮件签名与邮件加密的两个环节,而在接收端则相应增加邮件签名认证与邮件解密两个环节。邮件签名的作用是提供邮件完整性与不可抵赖服务,邮件加密的作用是提供邮件的保密服务。

数字信封技术可以用于电子邮件安全中。数字签名的作用能够保证邮件的完整性、身份认证与不可抵赖性,数据加密的作用可以保证邮件内容的机密性。

问题8-13:传输层安全协议SSL、PCT、TLS及OpenSSL之间是什么样的关系?

为了回答这个问题,需要注意以下几点。

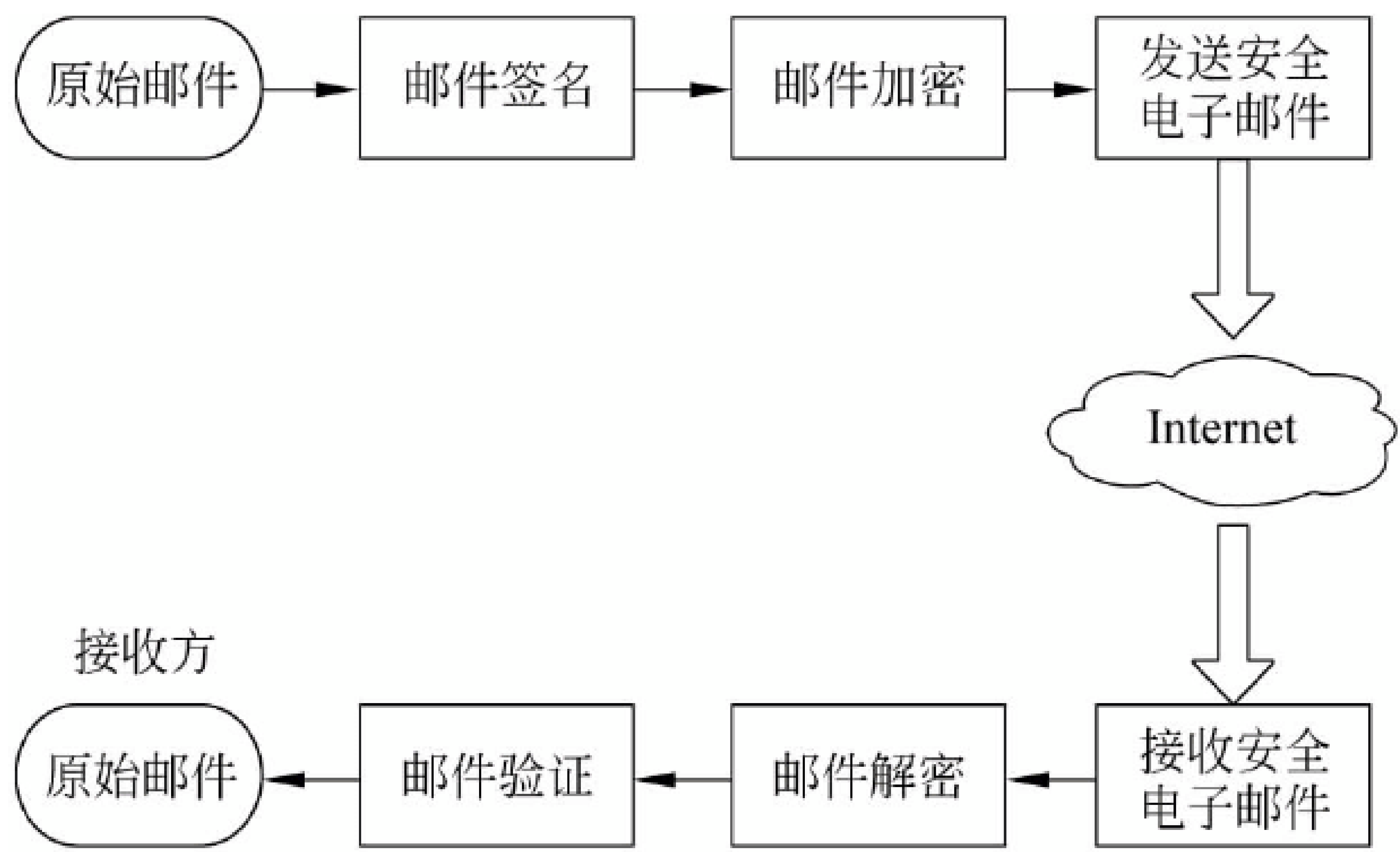


图 8-6　安全电子邮件的工作模式

（1）SSL、PCT 与 TLS 是三种传输层的安全协议，而 OpenSSL 是执行 SSLv3 协议的开源软件包。

（2）传输层安全协议的发展过程如下。

① 1994 年，Netscape 公司提出了用于 Web 应用的传输层安全协议 SSLv1，但是 SSLv1 协议没有实际使用过。1995 年，Netscape 开发了 SSLv2，并用于 Web 浏览器 Netscape Navigator 1.1 中。成为国际上最早应用于电子商务的一种网络安全协议。

② Microsoft 公司开发了类似的 PCT（Private Communication Technology）协议。Netscape 公司在 SSLv2 的基础上做了较大改进之后推出了 SSLv3。

③ IETF 鉴于 SSL 与 PCT 不兼容的现状，研发了 TLP（Transport Layer Protocol），希望推动传输层安全协议标准化。文档 RFC2246 对 TLP 进行了详细的描述。但是，SSL 与 PCT 也不完全兼容。

（3）目前，世界各国的网上支付系统广泛应用的仍然是 SSLv3 版协议。Eric A. Young 和 Tim J. Hudson 自 1995 年开始编写在 SSL 安全协议应用领域具有重大影响的开放源代码的 OpenSSL 软件包。1998 年，OpenSSL 项目组接管了 OpenSSL 的开发工作，并推出了 OpenSSL 的 0.9.1 版，到目前为止 OpenSSL 开源软件已经比较完善，支持 SSLv3 与 TLSv1 版本。

OpenSSL 采用 C 语言作为开发语言，方便了广大编程人员的使用；支持 Linux、Windows、BSD、Mac、VMS 等平台，这使得 OpenSSL 具有很好的跨平台性能。OpenSSL 软件由密码算法库、SSL 协议库以及应用程序等三个主要的功能部分组成，并提供了丰富的应用程序供测试功能。

问题 8-14：如何理解 SSL、SET 协议与 Web 安全的关系？

1. Web 安全问题的严重性

对 Web 系统安全构成威胁的因素很多。对于攻击者来说，Web 服务器、数据库服务器有很多的弱点可以被利用，比较明显的弱点在服务器的 CGI 程序与一些工具程序上。Web 服务的内容越丰富，应用程序越大，则包含错误代码的概率就越高。程序设计人员在编写 CGI 程序与一些工具程序时，一个简单的错误和不规范的编程都有可能形成系统的一个安全漏洞。对 Web 安全的研究一直是一个富有挑战性的课题。针对 Web 的攻击方法如图 8-7 所示。

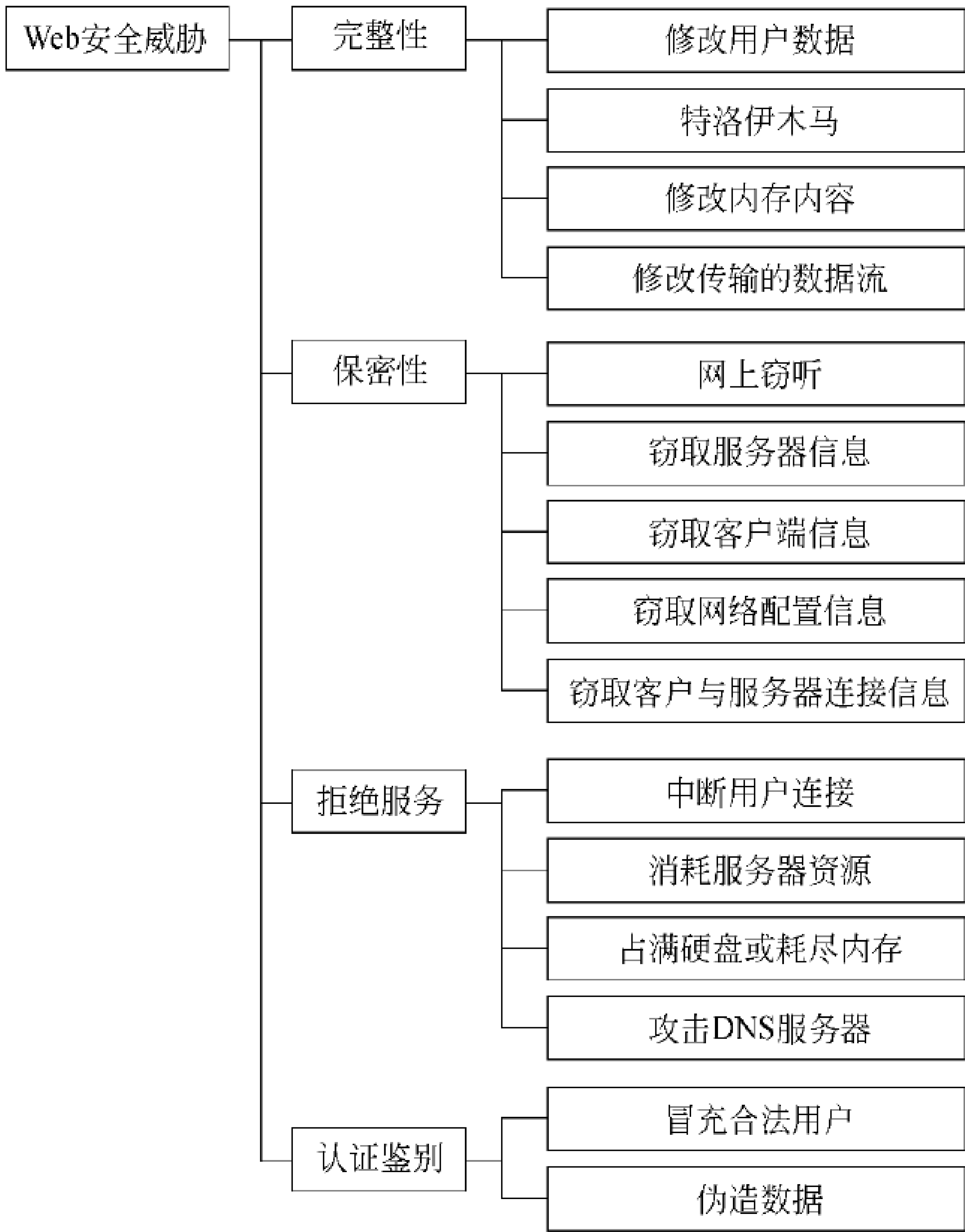


图 8-7　针对 Web 的攻击方法

2. Web 安全技术的研究思路

从网络体系结构的角度，Web 安全技术的研究可以分别从网络层、传输层和应用层协议着手。在网络层使用 IPSec 协议是设计安全 Web 体系首先会考虑的问题。在传输层可以采用安全套接层(Secure Sockets Layer，SSL)协议和传输层安全(Transport Layer Security，TLS)协议。在应用层 HTTP 之上采用安全电子交易(Secure Electronic Transaction，SET)协议。

3. SSL 协议

SSL 协议具有以下几个特点。

(1) SSL 协议尽管可以用于 HTTP、FTP、TELNET 等，但是目前主要应用于 HTTP，为基于 Web 服务的各种网络应用中的客户与服务器之间的用户身份认证与安全数据传输提供服务。

(2) SSL 协议处于端系统的应用层与传输层之间，在 TCP 之上建立一个加密的安全通道，为 TCP 之间传输的数据提供安全保障。

(3) 当 HTTP 使用 SSL 协议时，HTTP 的请求、应答报文格式与处理方法不变。它们的不同之处是：应用进程所产生的报文将通过 SSL 协议加密之后，再通过 TCP 连接传送出去；在接收端的 TCP 将加密的报文传送给 SSL 协议解密之后，再传送到应用层 HTTP。

(4) 当 Web 系统采用 SSL 协议时，Web 服务器的默认端口号从 80 变换为 443；Web 客

户端使用https取代常用的http。

(5) SSL协议包含两个协议：SSL握手协议(SSL Handshake Protocol)与SSL记录协议(SSL Record Protocol)。SSL握手协议实现双方加密算法的协商与密钥传递，SSL记录协议定义SSL数据传输格式，实现对数据的加密与解密操作。

4. SET协议

1) SET产生的背景

电子商务是以Internet环境为基础，在计算机系统支持下进行的商务活动。它是基于Web浏览/服务器应用方式，实现网上购物、网上交易和在线支付的一种新型商业运营模式。

2) SET提供的安全服务

基于Web的电子商务需要以下几个方面的安全服务。

(1) 鉴别贸易伙伴、持卡人的合法身份，以及交易商家身份的真实性。

(2) 确保订购与支付信息的保密性。

(3) 保证在交易过程中数据不被非法篡改或伪造，确保信息的完整性。

(4) 能够在TCP/IP之上运行，不抵制其他的安全协议的使用，不依赖特定的硬件平台、操作系统与Web软件。

SET协议是由VISA和MastCard两家信用卡公司与1997年提出的，并且已经成为目前公认的最成熟的应用层电子支付安全协议。SET协议使用了常规的对称加密与非对称加密体系，以及数字信封技术、数字签名技术、信息摘要技术与双重签名技术，以保证信息在Web环境中传输和处理的安全性。

3) SET系统结构

为了保证电子商务、网上购物与网上支付的安全性，SET协议定义了体系结构、电子支付协议与证书管理过程。

基于SET协议构成的电子商务系统由6个部分组成：持卡人、商家、发卡银行、收单银行、支付网关、认证中心。

SET结构的设计思想是：在持卡人、商家与收单银行之间建立一个可靠的金融信息传递关系，解决网上三方支付机制的安全性。

SET协议规定了加密算法的应用、证书授权过程与格式、信息交互过程与格式、认证信息格式等，使不同软件厂商开发的软件具有兼容性和互操作性，并能够运行在不同的硬件和操作系统平台上。

问题8-15：防火墙的设计思想是什么？

防火墙(Firewall)是在网络之间执行控制策略的系统，它包括硬件和软件。在设计防火墙时，人们做了一个假设：防火墙保护的内部网络是“可信赖的网络”，而外部网络是“不可信赖的网络”。设置防火墙的目的是保护内部网络资源不被外部非授权用户使用，防止内部受到外部非法用户的攻击。那么防火墙安装的位置一定是在内部网络与外部网络之间。防火墙的主要功能如下。

(1) 检查所有从外部网络进入内部网络的数据包。

(2) 检查所有从内部网络流出到外部网络的数据包。

(3) 执行安全策略,限制所有不符合安全策略要求的分组通过。

(4) 具有防攻击能力,保证自身的安全性。

问题8-16:如何认识防火墙系统的结构?

实际上由于网络的安全策略与防护要求不同,防火墙系统的配置与结构也有很大的区别,可以采用多级的结构。任何一种结构的防火墙系统都是由包过滤路由器与应用级网关组合而成。为了使防火墙系统结构配置的表示更简洁、明确,可以引入一些符号。例如,S表示包过滤路由器,B1表示只有一个网络接口的堡垒主机,B2表示有两个网络接口的堡垒主机。

1. 安全缓冲区或非军事区的概念

在设计复杂结构的防火墙系统时,人们的分析与设计思路是:将防火墙系统分为内外两个部分。从外包过滤路由器开始的部分是由网络系统所属的单位组建的,因此属于单位的内部网络。外包过滤路由器与外堡垒主机构成了防火墙的过滤子网。人们通常将必须向外部提供服务,并且安全要求相对比较低的服务器(如E-mail服务器)连接在过滤子网,而将安全要求比较高的服务器、工作站连接在内部网络的服务子网中。因此人们也把过滤子网叫作安全缓冲区或非军事区。内包过滤路由器与内堡垒主机用来进一步保护内部网络的服务器与工作站。如果外部用户希望访问内部网络的服务,那么它需要经过两级过滤路由器与堡垒主机的审查,因此外部非法用户进入系统内部网络成功的可能性将会大大降低,但是这是以高造价和访问速度的降低为代价的。

2. 防火墙系统结构

教材中对比较典型的采用一个过滤路由器与一个堡垒主机组成的S-B1防火墙系统结构、采用两个过滤路由器与两个堡垒主机组成的S-B1-S-B1配置的防火墙系统结构做了详细的分析。实际上防火墙的系统结构是多样的。

1) S-B1防火墙系统的两种构成方式

图8-8给出了由一个过滤路由器与一个堡垒主机组成的S-B1防火墙系统结构,还可以有如图8-9所示的S-B2结构形式。在S-B1防火墙系统结构中,堡垒主机B1是接在过滤子网与内部网络之间的。尽管它们的结构不同,但是从网络的层次结构与数据在防火墙系统中的传输、处理过程都可以表示为如图8-10所示的形式。

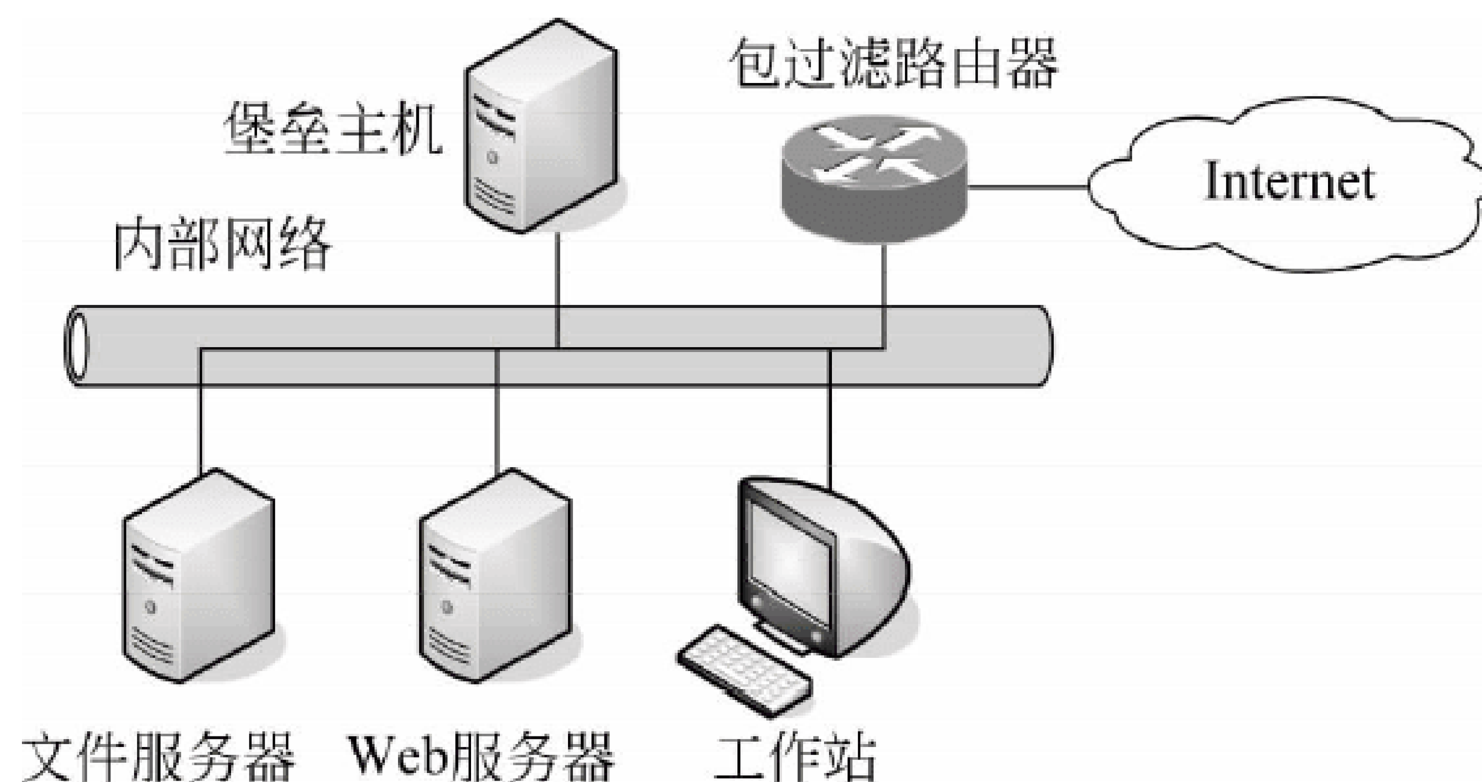


图8-8 S-B1防火墙系统结构

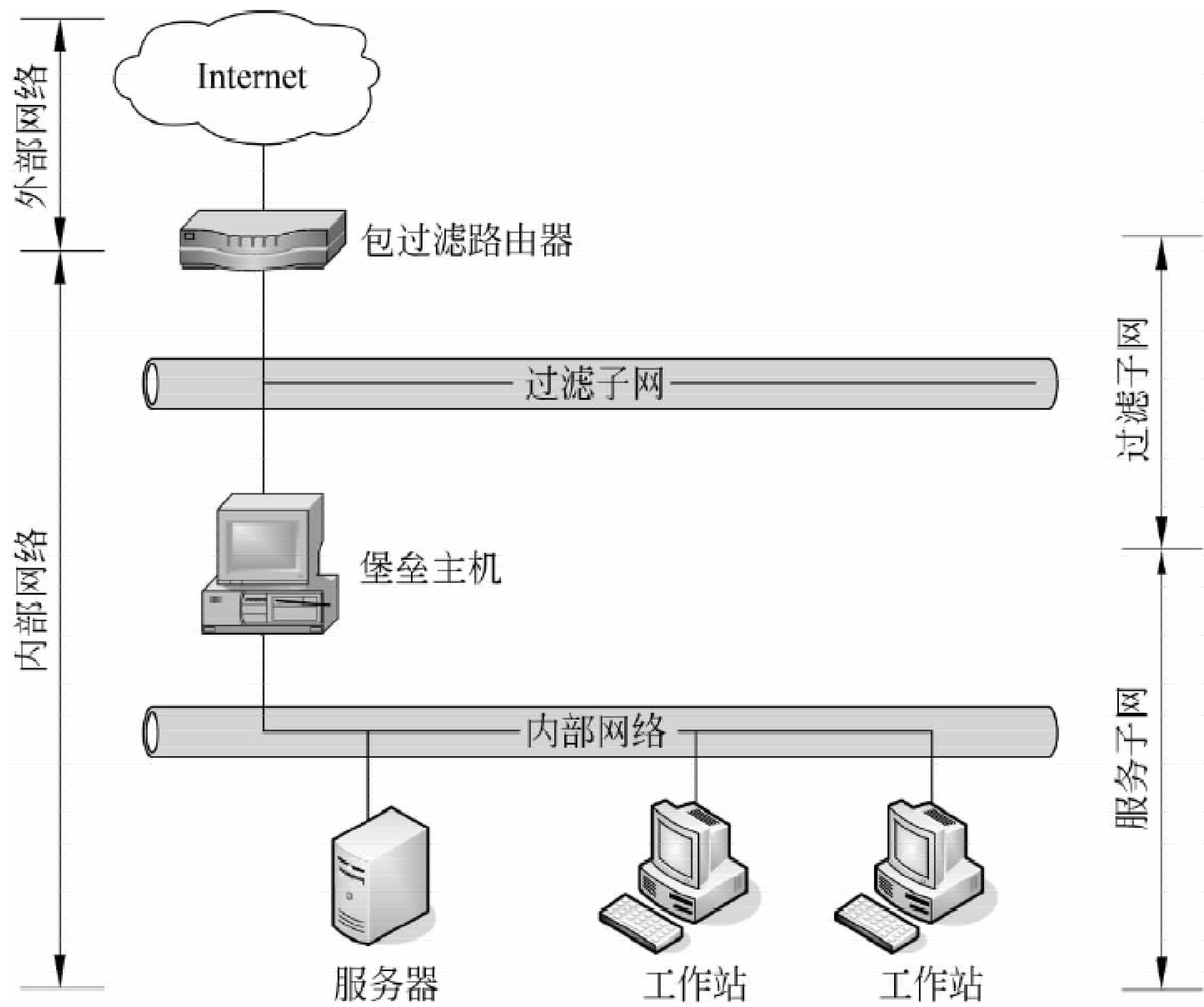


图 8-9　S-B2 防火墙系统结构

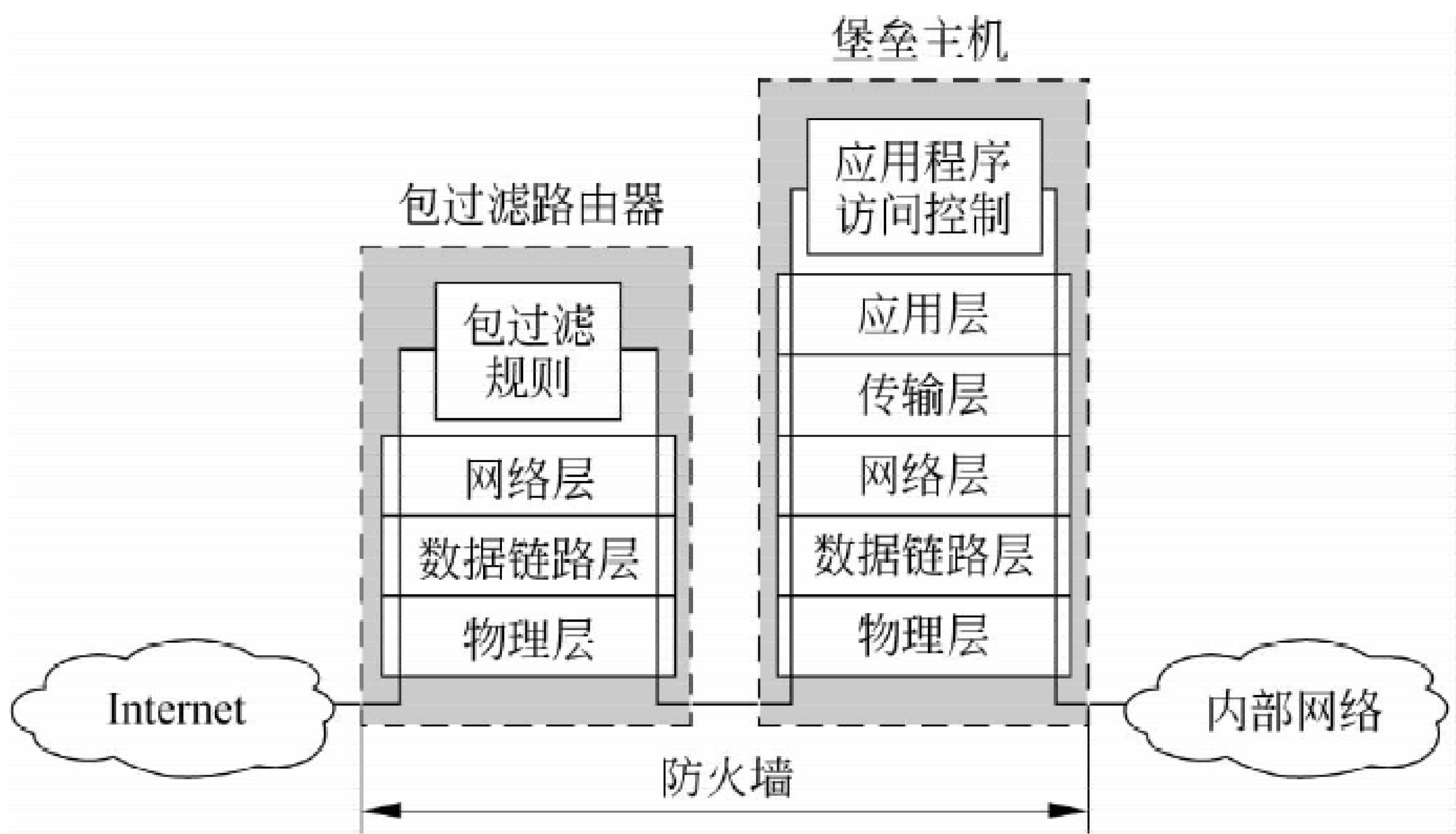


图 8-10　S-B1 与 S-B2 配置的防火墙系统层次结构示意图

2）S-B2-B2 防火墙系统结构

图 8-11 给出了 S-B2-B2 防火墙系统结构。图 8-12 给出了 S-B2-B2 配置的防火墙系统结构中的数据处理与传输过程示意图。

3）S-B1-S-B1 防火墙系统结构

图 8-13 给出了 S-B1-S-B1 防火墙系统结构。图 8-14 给出 S-B1-S-B1 配置的防火墙系统结构中的数据处理与传输过程示意图。

我们也将过滤子网叫作安全缓冲区或非军事区(DMZ)。DMZ 是指一个非保护的网络区域,任何非敏感、能够被外部用户直接访问的服务器(例如 Web、E-Mail 服务器)可以放置在 DMZ 中。DMZ 中的系统可能是最先受到攻击和被破坏的系统,但是它的状态对于内部网络系统的安全不会造成影响。

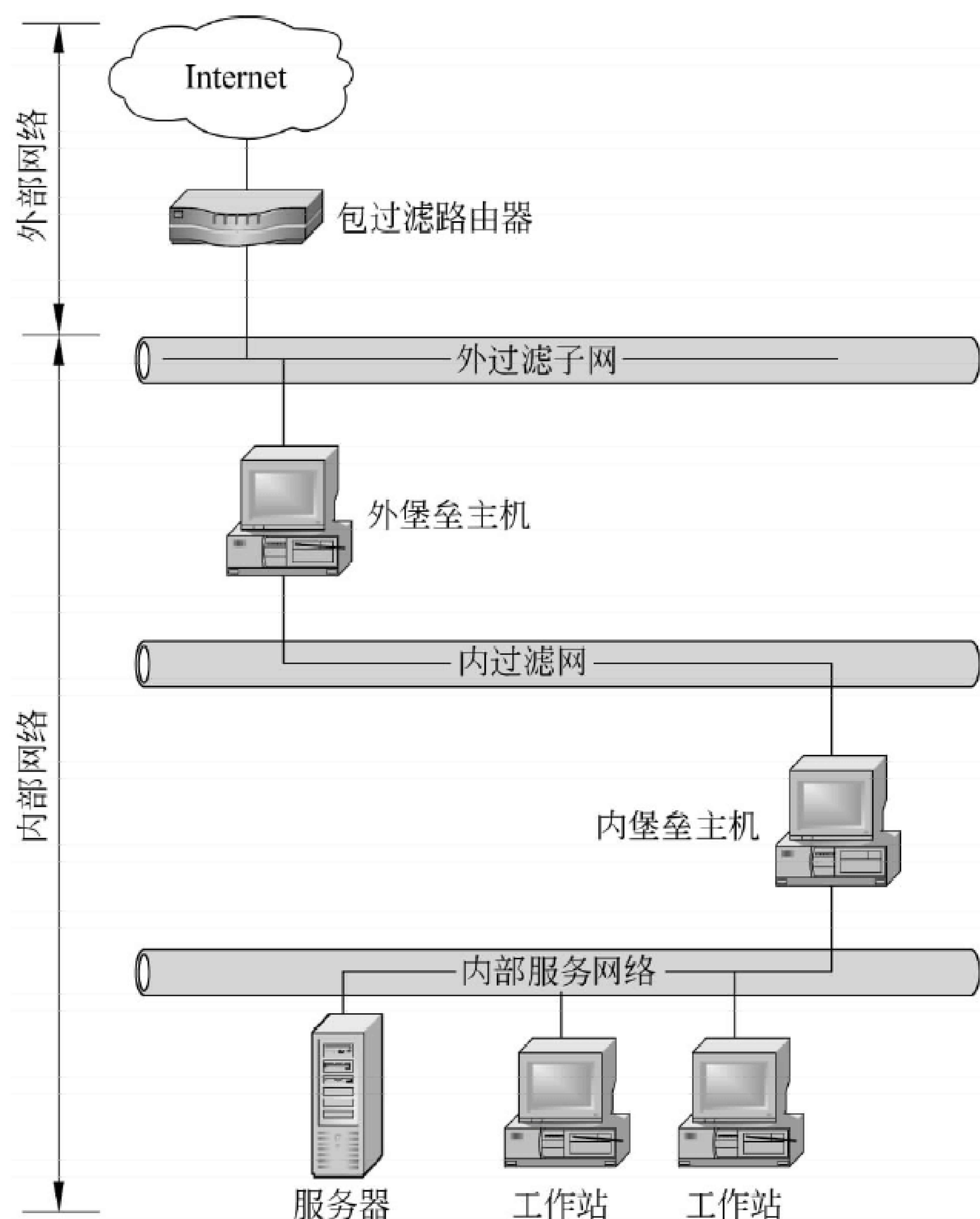


图 8-11　S-B2-B2 防火墙系统结构

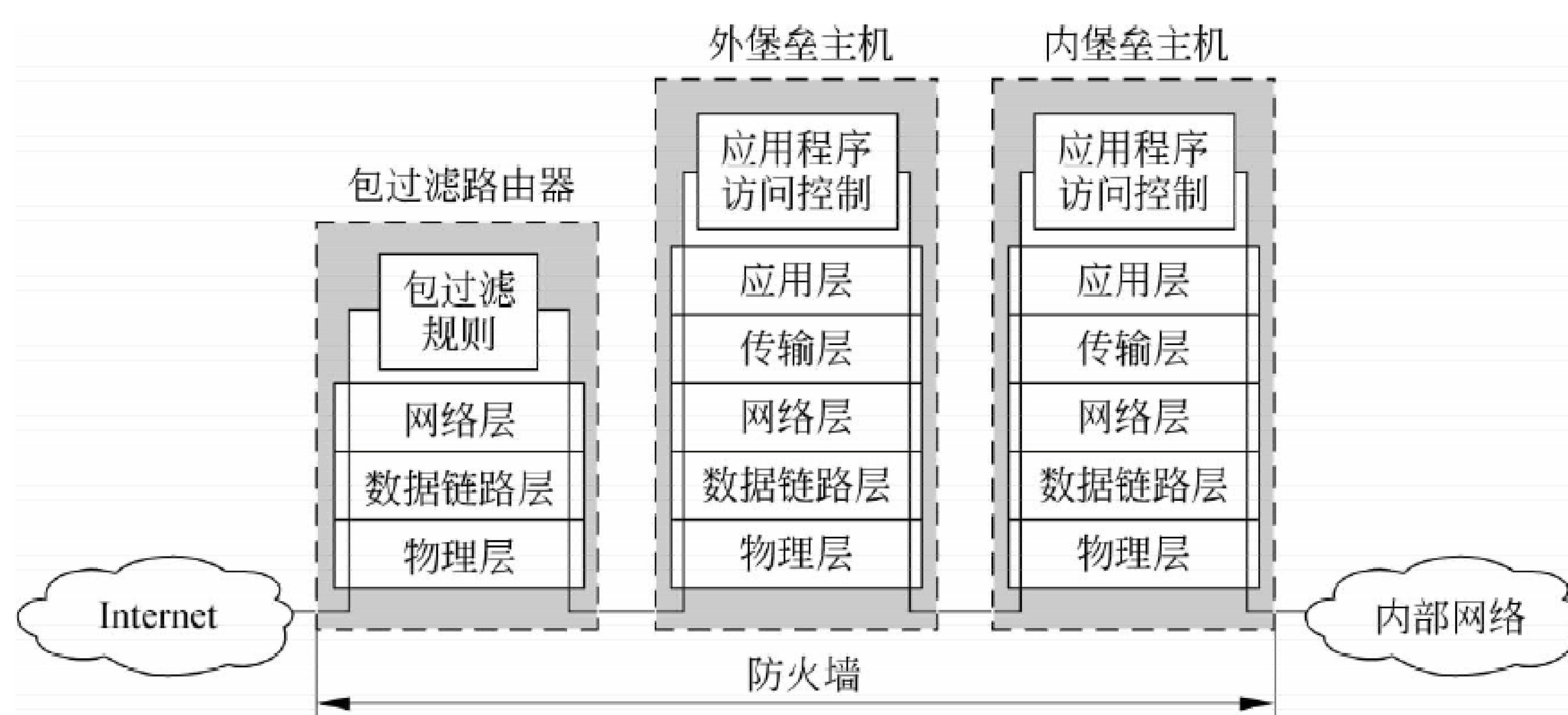


图 8-12　S-B2-B2 配置的防火墙系统结构层次结构示意图

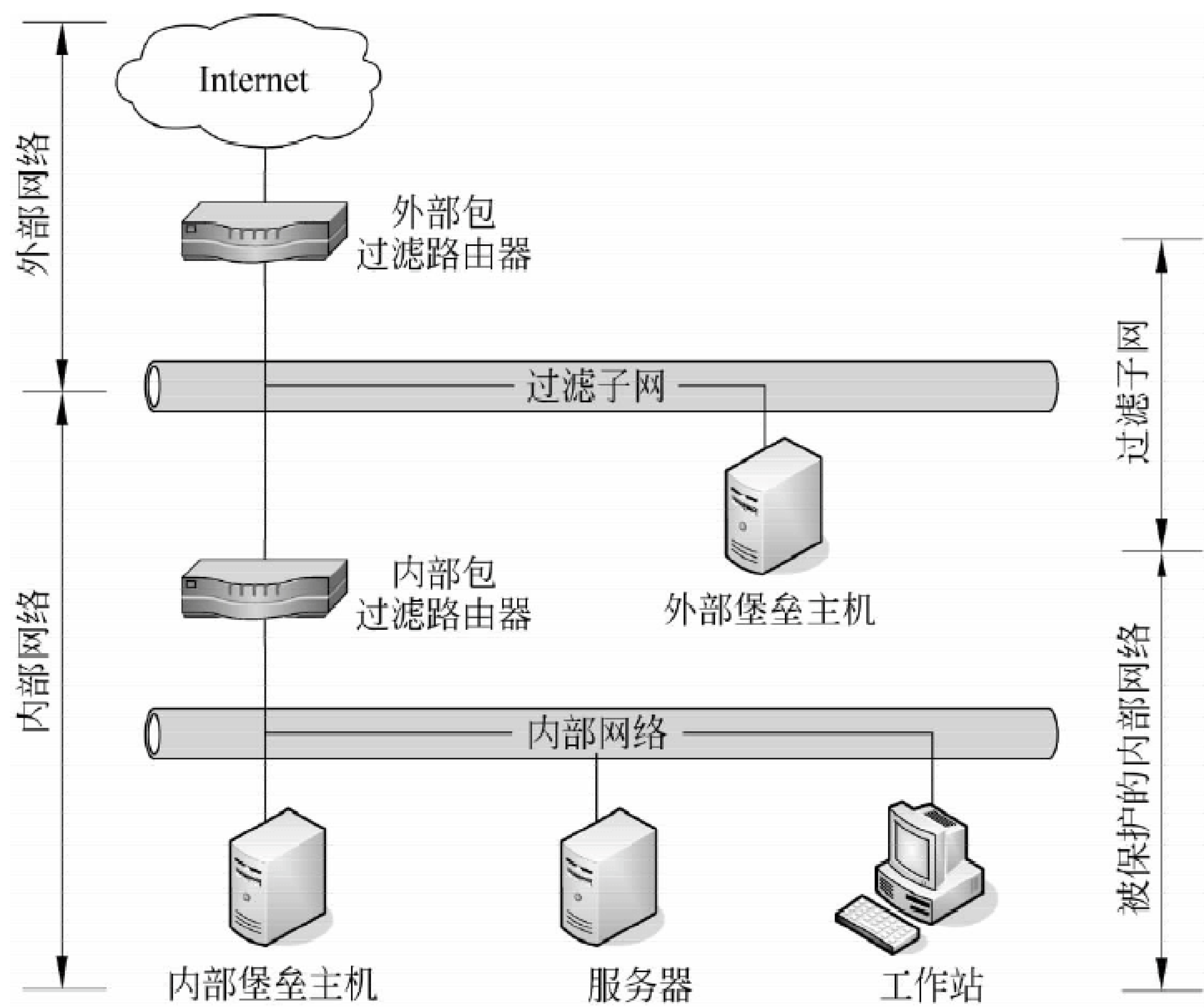


图 8-13　S-B1-S-B1 防火墙系统结构

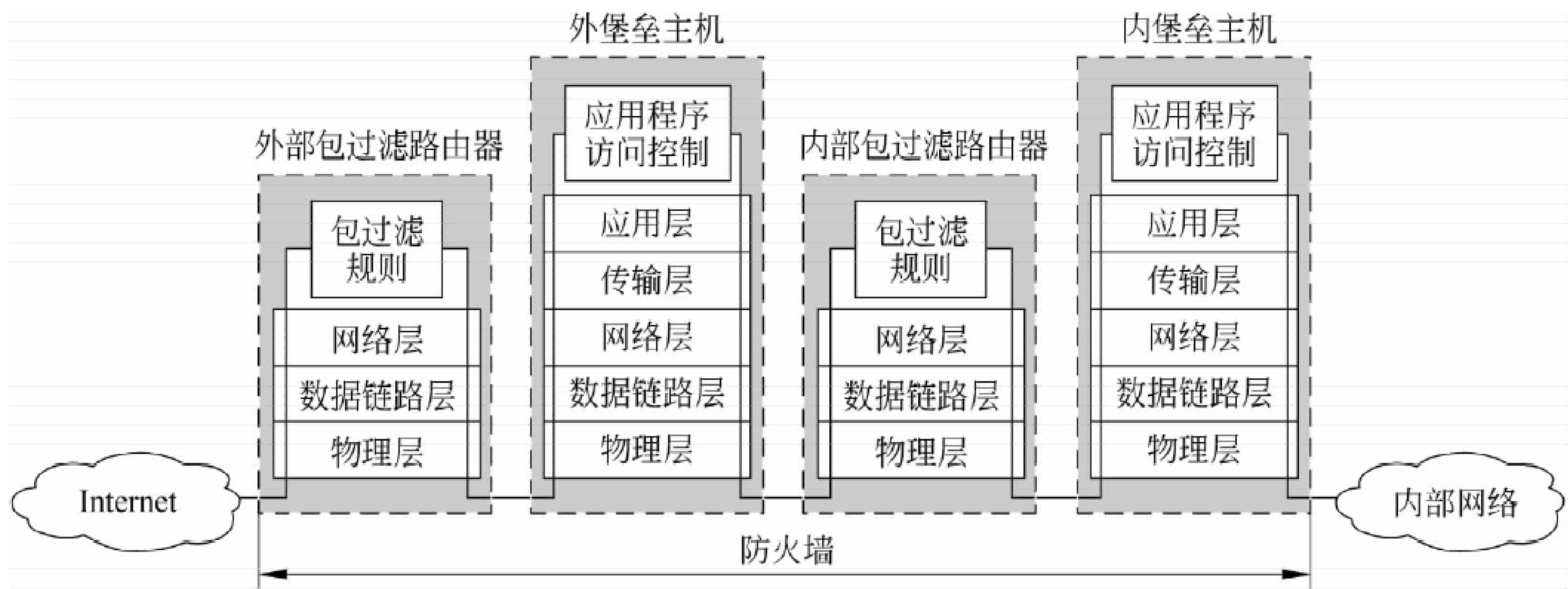


图 8-14　S-B1-S-B1 配置的防火墙系统层次结构示意图

问题 8-17：防火墙技术有什么样的局限性?

防火墙的主要功能就是控制对受保护网络的非法访问。防火墙通过监视、限制、更改通过防火墙的数据包，一方面对外部用户屏蔽内部网络的拓扑结构，限制外部非法用户访问内部；另一方面限制内部用户访问外部危险站点。但是，防火墙也有其明显的局限性，这主要表现在以下几点。

1. 防火墙系统对内部的防护能力较弱

防火墙系统的安全控制功能是有限的，它主要是用于对外屏蔽内部网的拓扑结构，控制外部网上的用户访问内部网上的重要站点或某些端口以及对内屏蔽外部危险站点，但它很难解决内部人员违反网络使用规定所引起的安全问题。统计结果表明，网络上的安全攻击事件有 70%以上来自内部攻击。

2. 防火墙系统难以配置和管理,容易造成安全漏洞

防火墙系统的配置与管理相当复杂,要想成功地维护防火墙,管理员对网络安全攻击的手段与系统配置的关系必须有相当深刻的了解。由多个路由器、包过滤路由器、应用网关组成的防火墙系统,管理上稍有疏忽就有可能造成潜在的危险。统计表明,30%的入侵是在有防火墙的情况下发生的。

3. 防火墙系统难以为不同用户提供不同的安全控制策略

许多简单的防火墙的安全控制主要是基于用户所用机器的 IP 地址,而不是用户身份。这样就很难为同一台主机的不同用户提供不同的安全控制策略。

无论什么事都有它的两面性。在看到防火墙在安全防范中的积极作用的同时,也要看到防火墙的局限性,并且应该清醒地认识到有时防火墙会给人一种不实际的安全感,导致内部管理的松懈,很多内部的攻击行为不是任何基于隔离作用的防火墙所能防范的。因此,构筑网络系统的安全体系,必须将防火墙和其他技术手段以及网络管理统一起来考虑。

问题 8-18:网络攻击包括哪些主要的类型?

回答这个问题需要注意以下几点。

1. 网络攻击的类型

目前,网络攻击大致可以分为系统入侵类攻击、缓冲区溢出攻击、欺骗类攻击与拒绝服务 DoS 攻击等。系统入侵类攻击者的最终目的都是为了获得主机系统的控制权,从而破坏主机和网络系统。这类攻击又分为信息收集攻击、口令攻击、漏洞攻击。缓冲区溢出攻击是指:通过往程序的缓冲区写超出其长度的内容,造成缓冲区的溢出,从而破坏程序的堆栈,使程序转而执行其他指令。缓冲区攻击的目的在于扰乱那些以特权身份运行的程序的功能,使攻击者获得程序的控制权。网络欺骗的主要类型有 IP 欺骗、ARP 欺骗、DNS 欺骗、Web 欺骗、电子邮件欺骗、源路由欺骗、地址欺骗与口令欺骗等。

2. 网络安全威胁的层次

网络安全威胁可以分为三个层次:主干网络的威胁、TCP/IP 安全威胁与网络应用的威胁。主干网络的威胁主要表现在主干路由器与 DNS 服务器。攻击主干网最直接的方法就是攻击它的主干路由器与 DNS 服务器。2002 年 8 月,黑客利用 Internet 主干网的 ASN No.1 信令存在的安全漏洞,攻击了主干路由器、交换机和一些基础设施,造成了严重的后果。全球 13 台根域 DNS 服务器支持着整个 Internet 的运行。1997 年 7 月的人为错误曾经导致根域 DNS 服务器工作不正常,致使 Internet 系统局部服务中断。2002 年 10 月 21 日美国东部时间下午 4:45 开始,13 台根域 DNS 服务器遭受了规模最大的分布式拒绝服务 DDoS 攻击,导致其中的 9 台根域 DNS 服务器工作不正常。2016 年 10 月 21 日,为美国众多公司提供互联网网络域名解析服务的美国 DNS 域名服务提供商 Dyn 位于美国东海岸 DNS 基础设施,受到来自全球的 DDoS 攻击,攻击事件一直持续到当天 13:45,攻击事件涉及的 IP 地址的数量达到千万量级,其中很大部分来自物联网与智能系统。被百度感染的网络设备包括路由器、网络摄像头、DVR 设备等。受到攻击影响的厂商包括 Twitter、Etsy 等。这次攻击暴露出物联网与智能系统的很多安全隐患。

3. 服务攻击与非服务攻击的基本概念

Internet 中的网络攻击可以归纳为两种基本类型:服务攻击与非服务攻击。

服务攻击是指对为网络提供某种服务的服务器发起攻击,造成该网络的“拒绝服务”,使

网络工作不正常。特定的网络服务包括 E-mail、Telnet、FTP、WWW 服务等。

非服务攻击不针对某项具体应用服务，而是针对网络层及低层协议进行的。攻击者可能使用各种方法对网络通信设备(例如路由器、交换机)发起攻击，使网络通信设备工作严重阻塞或瘫痪。

4. 网络攻击手段的分类

网络攻击手段的分类如图 8-15 所示。

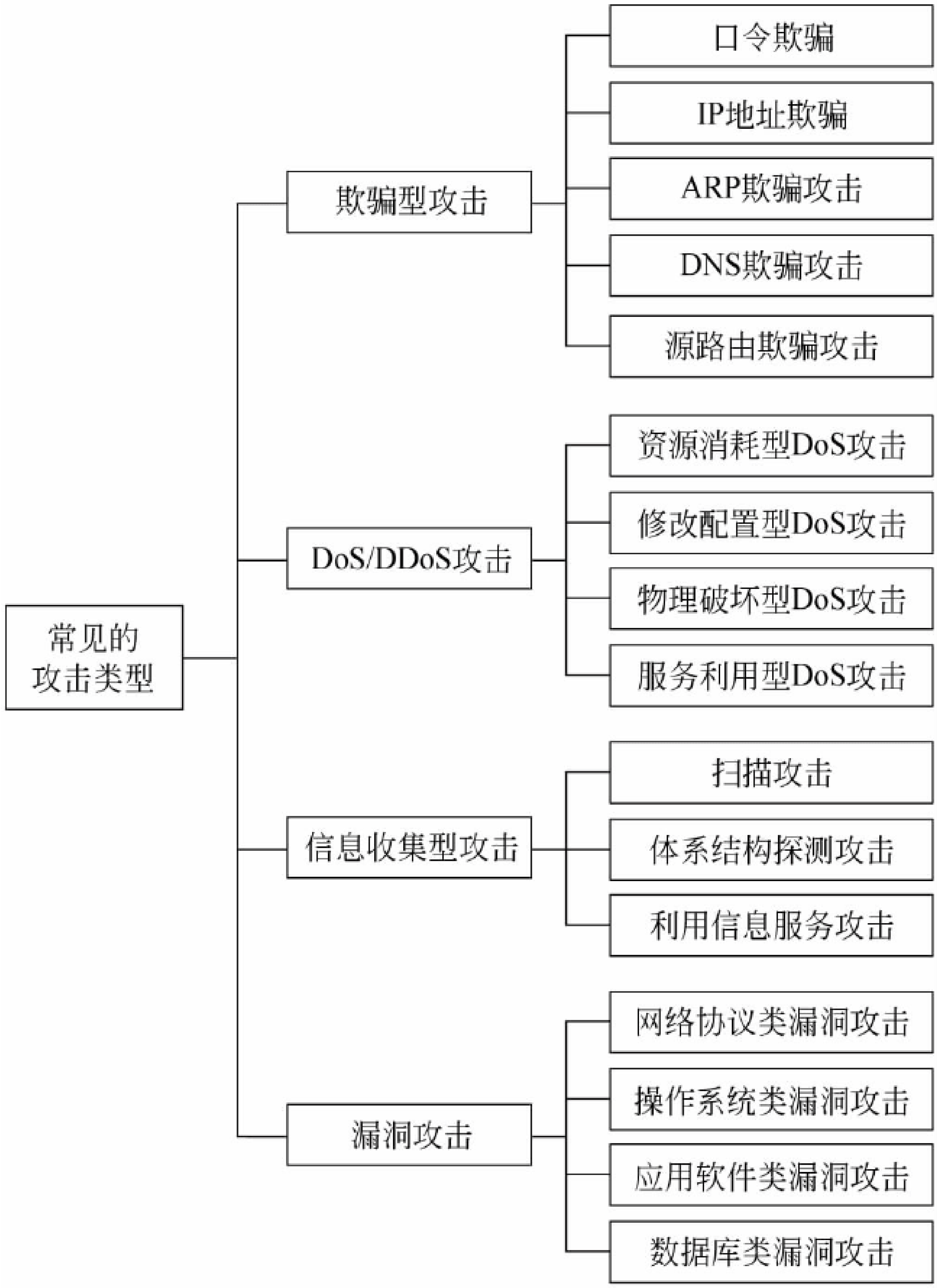


图 8-15　网络攻击手段的分类

网络攻击手段很多并且不断地变化，总结目前出现的主要的网络攻击现象与手段大致可以将它们分为欺骗类攻击、DoS/DDoS 类攻击、信息收集类攻击、漏洞类攻击等 4 种基本类型。

1) 欺骗类攻击

欺骗类攻击的手段主要包括口令欺骗、IP 地址欺骗、ARP 欺骗、DNS 欺骗与源路由欺骗等。

2) DoS/DDoS 攻击

拒绝服务 DoS 攻击与分布式拒绝服务 DDoS 攻击的手段主要包括资源消耗型、修改配置型、物理破坏型与服务利用型等类型的拒绝服务攻击。

3) 信息收集类攻击

信息收集类攻击的手段主要包括：扫描攻击、体系结构探测攻击、利用服务攻击等。

4) 漏洞类攻击

漏洞类攻击的手段主要包括网络协议类、操作系统类、应用软件类与数据库类等。

同时，需要注意的是，网络安全漏洞实际上分为两大类：技术漏洞与管理漏洞等。这里主要考虑的是技术漏洞类的问题。

问题8-19：如何认识DoS攻击与DDoS攻击？

1. 拒绝服务DoS攻击

拒绝服务(DoS)攻击主要是通过消耗网络系统有限的、不可恢复的资源，从而使合法用户应该获得的服务质量下降或受到拒绝。DoS攻击最本质的是延长正常网络应用服务的等待时间，或者使合法用户的服务请求受到拒绝。DoS攻击的目的不是闯入一个站点或者是更改数据，而是使站点无法服务于合法的服务请求。拒绝服务DoS攻击的分类如图8-16所示。

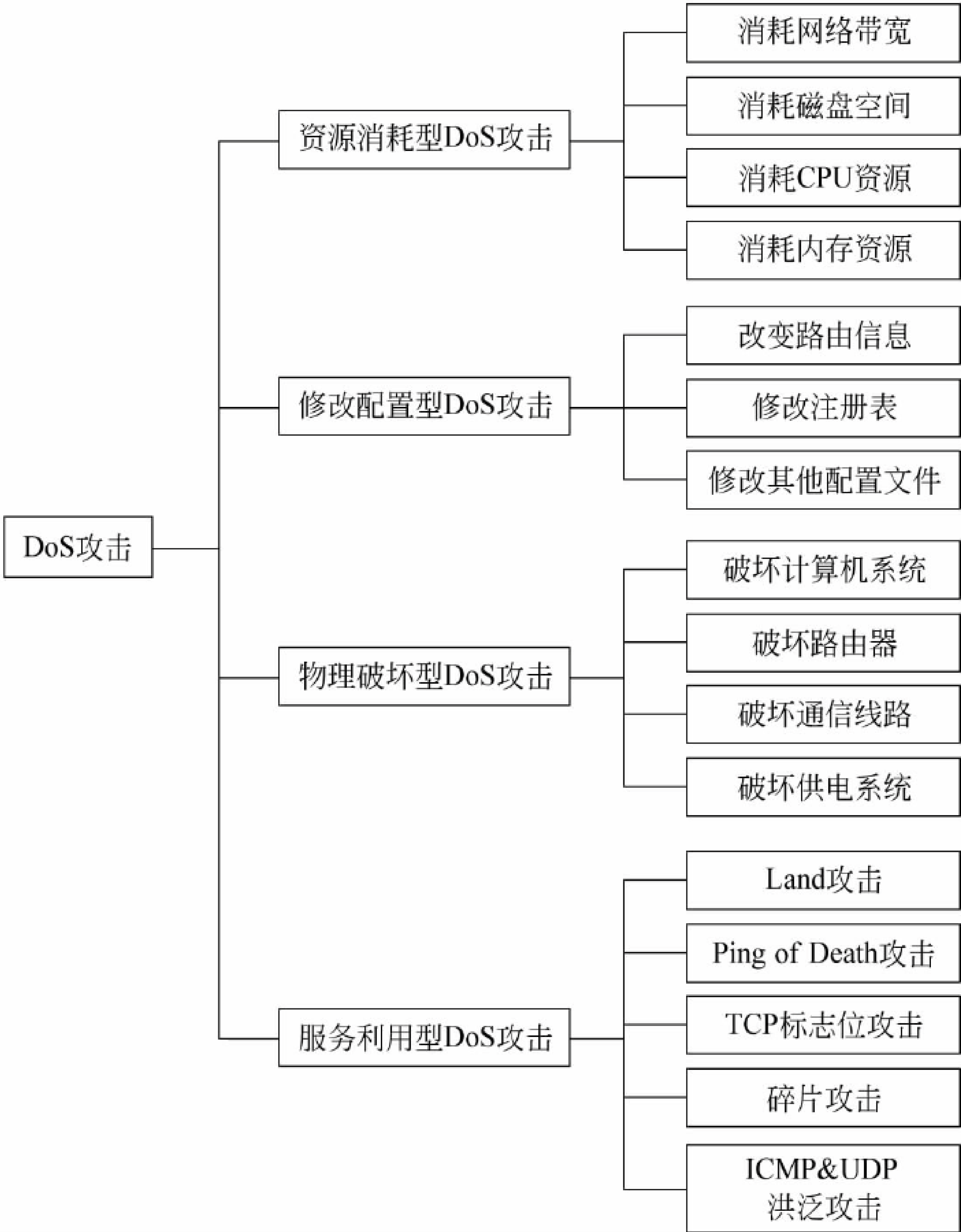


图8-16　DoS攻击的分类

拒绝服务DoS攻击大致可以分为4类：资源消耗型DoS攻击、修改配置型DoS攻击、物理破坏型DoS攻击、服务利用型DoS攻击。

1）资源消耗型 DoS 攻击

资源消耗型 DoS 攻击通过消耗网络带宽、内存和磁盘空间、CPU 利用率，使网络系统不能正常工作。常见的方法如下。

（1）攻击者制造大量广播包或传输大量文件，占用网络链路与路由器带宽资源。

（2）攻击者制造大量电子邮件、错误日志信息、垃圾邮件等，占用主机中共享的磁盘资源。

（3）攻击者制造大量无用信息或进程通信交互信息，占用 CPU 和系统内存资源。

2）修改配置型 DoS 攻击

修改配置型 DoS 攻击通过修改系统运行配置，阻止合法用户的使用和网络的正常工作。常见的方法如下。

（1）改变路由信息。

（2）修改 Windows NT 的注册表。

（3）修改 UNIX 的各种配置文件。

3）物理破坏型 DoS 攻击

物理破坏型 DoS 攻击通过破坏网络、计算机或系统物理支持环境，使网络系统不能正常工作。常见的方法如下。

（1）破坏计算机系统。

（2）破坏路由器和通信线路。

（3）破坏网络与计算机设备供电或机房空调系统。

4）服务利用型 DoS 攻击

服务利用型 DoS 攻击利用网络或协议的漏洞达到攻击的目的。常见的方法如 Land 攻击、Ping to Death 攻击、TCP 标志位攻击、IP 碎片攻击、ICMP& UDP 洪泛攻击等。

2. 分布式拒绝服务

分布式拒绝服务(DDoS)攻击是在 DoS 攻击基础上产生的一类攻击形式。DDoS 攻击采用了一种比较特殊的体系结构，攻击者利用多台分布在不同位置的攻击代理主机，同时攻击一个目标，从而导致被攻击者的系统瘫痪，其攻击过程如图 8-17 所示。

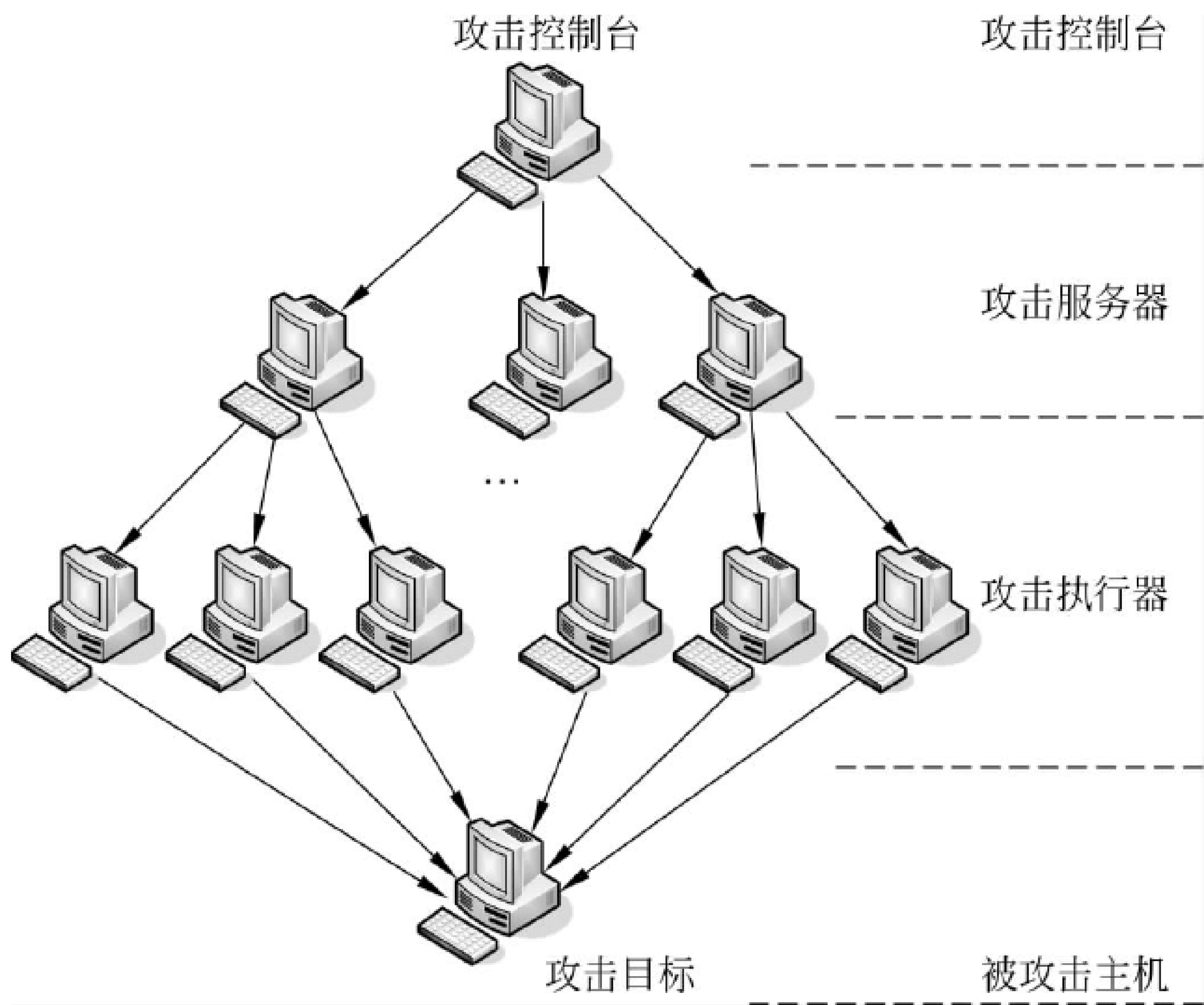


图 8-17　DDoS 攻击过程示意图

典型的 DDoS 攻击采用的是三层结构：攻击控制层、攻击服务器层、攻击执行器层。DDoS 攻击是建立在许多与攻击无关的主机被动地被支配的前提之下实现的。攻击控制台可以是网络上的任何一台主机，甚至是一台移动的便携机，它的作用是向攻击服务器分布攻击命令。攻击服务器的主要任务是将攻击控制台的命令分布到攻击执行器。攻击服务器与攻击执行器都已经被攻击控制器侵入，并被暗地里安装了攻击软件。

DDoS 攻击的第一步是攻击者选择一些防护能力弱的主机或服务器，通过寻找系统漏洞或系统配置错误，成功侵入并安装后门程序。有时攻击者也需要进一步通过网络监听，增加被侵入的主机数量。

第二步是在入侵的主机系统中安装攻击服务器软件或攻击执行器软件。攻击服务器数量一般在几台到几十台。设置攻击服务器的目的是隔离网络的联系渠道，防止被追踪，保护攻击者。攻击执行器安装相对简单的攻击软件，它只需要连续向攻击目标主机发送大量的连接请求，而不做任何应答。

第三步就是攻击控制台向攻击服务器发出攻击命令，由多个攻击服务器再向攻击执行器发出攻击命令，攻击执行器同时向目标主机发起攻击。在向攻击服务器发出攻击命令的很短的时间内，攻击控制台可以立即撤离网络，使得追踪很难实现。

DDoS 攻击的特征如下。

(1) 被攻击主机上有大量等待的 TCP 连接。

(2) 网络中充斥着大量的无用数据包，并且数据包的源地址是伪造的。

(3) 大量无用数据包造成网络拥塞，使被攻击的主机无法正常地与外界通信。

(4) 被攻击主机无法正常回复合法用户的服务请求。

(5) 严重时会造成主机系统瘫痪。

目前，典型的 DDoS 攻击软件有 Trinoo、Tribe Flood Network(TFN)、Tribe Flood Network 2000(TFN2K)、Stacheldraht、Shaft 与 Mstream 等。

问题 8-20：如何认识僵尸网络的特征？

为了认识僵尸网络的特征，需要注意以下几个问题。

(1) 僵尸网络是在蠕虫、特洛伊木马、后门等恶意代码基础上发展出来的一种新型攻击方式。从 1999 年第一个具有僵尸网络特性的恶意代码 PrettyPark 出现，到 2002 年因 SDbot 和 Agobot 的广泛传播，僵尸网络对 Internet 构成了严重的安全威胁。

(2) 反病毒厂商一直没有给出僵尸网络和僵尸程序的准确定义，仍然将它归于蠕虫或后门的范畴。从 2003 年前后，学术界开始关注这种新兴的安全威胁，并将僵尸网络定义为：僵尸网络是控制者出于恶意目的，传播僵尸程序控制大量主机，并通过一对多的命令与控制信道所组成的网络。

问题 8-21：如何认识入侵检测技术的特征？

入侵检测系统(Intrusion Detection System，IDS)是对计算机和网络资源的恶意使用行为进行识别的系统。它的目的是监测和发现可能存在的攻击行为，包括来自系统外部的入侵行为和来自内部用户的非授权行为，并采取相应的防护手段。

1. 入侵检测系统的基本功能

1980 年，James Anderson 在 *Computer Security Threat Monitoring and Surveillance*

的论文中提出了入侵检测系统的概念。1987 年，Domthy Donning 在论文 *An Intrusion Detection Model* 中提出了入侵检测系统 IDS 的框架结构。入侵检测系统的基本功能主要有以下几个。

(1) 监控、分析用户和系统的行为。

(2) 检查系统的配置和漏洞。

(3) 评估重要的系统和数据文件的完整性。

(4) 对异常行为的统计分析，识别攻击类型，并向网络管理人员报警。

(5) 对操作系统进行审计、跟踪管理，识别违反授权的用户活动。

2. 入侵检测系统的结构

图 8-18 给出了入侵检测系统 IDS 的通用框架结构(Common Intrusion Detection Framework，CIDF)。入侵检测系统一般是由事件发生器、事件分析器、响应单元与事件数据库组成。

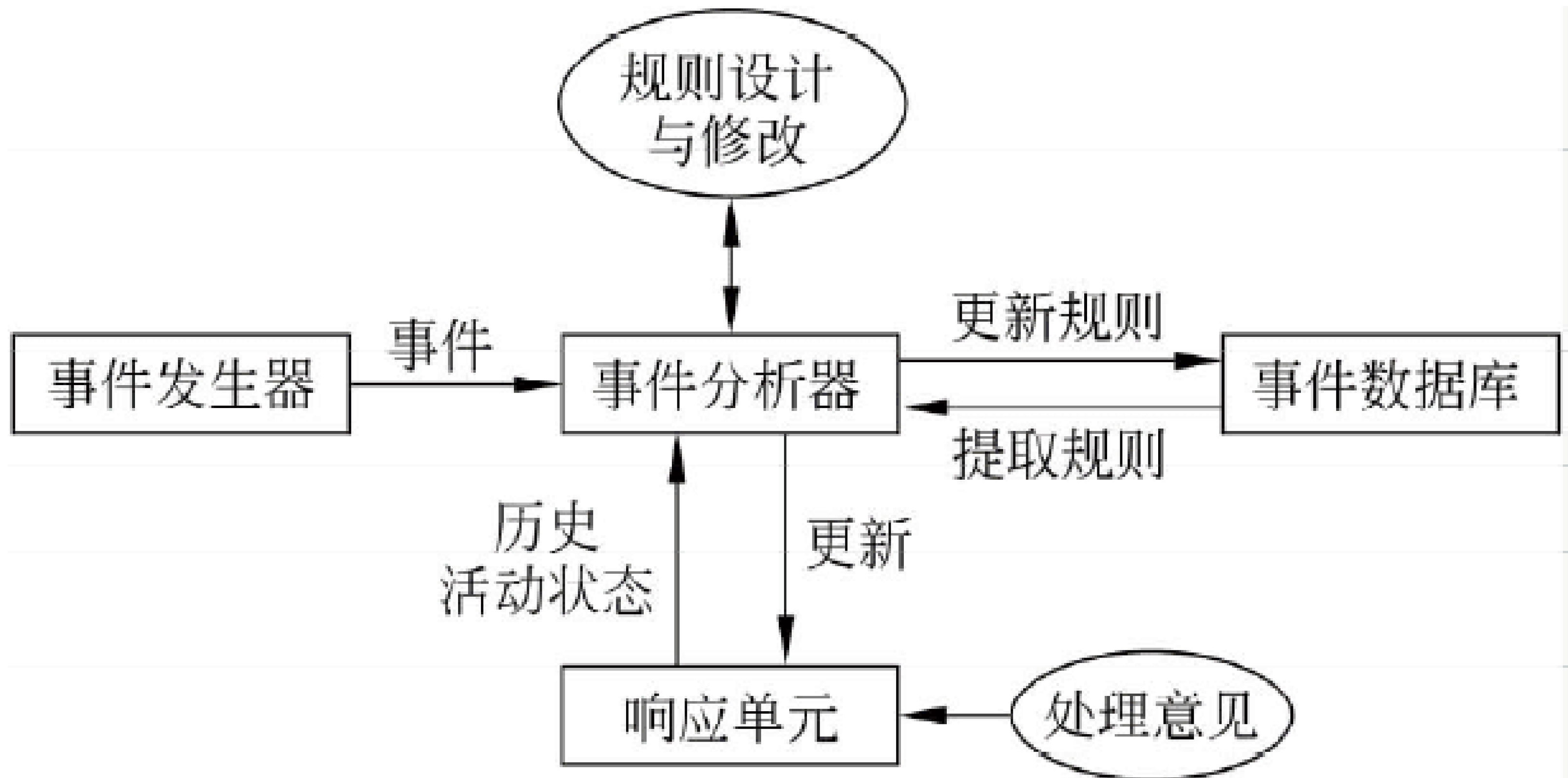


图 8-18 入侵检测系统 IDS 的通用框架结构

1) 事件发生器

CIDF 通用框架结构将入侵检测系统 IDS 需要分析的数据统称为事件(Event)，它可以是网络中的数据包，也可以是从系统日志等其他途径得到的信息。事件发生器产生的事件可能是经过协议解析的数据包，或者是从日志文件中提取的相关部分。

2) 事件分析器

事件分析器根据事件数据库的入侵特征描述、用户历史行为模型等，解析事件发生器产生的事件，得到格式化的描述，判断什么是合法的，什么是非法的。

3) 响应单元

响应单元则是对分析结果做出反应的功能单元，它可以做出切断连接、改变文件属性或报警等响应。

4) 事件数据库

事件数据库存放攻击类型数据或者检测规则，它可以是复杂的数据库，也可以是简单的文本文件。事件数据库储存有入侵特征描述、用户历史行为等模型和专家经验。

问题 8-22：恶意软件与病毒是什么关系?

恶意代码与病毒是有区别的。但是由于早期恶意代码的主要形式是病毒，用户平时都要接触病毒防治软件，因此人们习惯于用病毒去称呼恶意代码。回答恶意代码与病毒的区别，需要注意以下几点。

1. 从恶意代码的特征来看

恶意代码具有如下三个共同的特征：恶意的目的、本身是程序、通过执行发生作用。这一点显示出它是对一类恶意软件的统称。

2. 从恶意代码的发展阶段来看

从1986年第一个可以自我复制的计算机病毒Brain出现至今，恶意代码大致经历了4个发展阶段：DOS病毒、宏病毒、网络蠕虫病毒与趋利性恶意代码。

尤其是从第三代、第四代恶意代码的产生充分地表现出趋利性的动机，使得恶意代码的类型发生了很大变化，目前的恶意代码已经包括病毒、特洛伊木马、蠕虫、垃圾邮件、流氓软件等多种形式。因此，从发展的角度看，病毒只是恶意代码中的一种。

3. 从各种恶意代码特征来看

从工作机理、传播途径、产生的后果来看，病毒、特洛伊木马、蠕虫、垃圾邮件、流氓软件之间是有区别的。但是病毒、特洛伊木马、蠕虫、垃圾邮件、流氓软件之间，以及与网络攻击之间呈现出融合的趋势，变种速度快，检测难度增加。例如，从用户的角度看垃圾邮件应该不属于病毒的范畴，但是很多垃圾邮件以及成为传播病毒的主要手段。早期的病毒与网络攻击是有区别的，但是现在的病毒已经带上了明显的网络攻击色彩与效果。

4. 从恶意代码的传播途径来看

恶意代码的传播途径主要是：利用操作系统或应用软件的漏洞、通过浏览器、利用用户的信任关系。这一点与网络攻击是很相似的。

问题8-23：如何理解恶意代码定义中“故意”的含义？

恶意代码的定义是：能够从一台计算机传播到另一台计算机，从一个网络传播到一个网络的程序，目的是在用户和网络管理员不知情的情况下对系统进行“故意”修改。这里突出了“故意”的特征。原因是：从法律的角度看，如果一个操作系统的某个版本或某种应用软件有重大的缺陷，它也可能造成系统的瘫痪、数据被删除、配置被修改。但是，从法理上说，它不是主观故意，软件公司需要承担责任，但是这个软件不属于恶意代码或恶意软件的范畴。因此，在恶意代码的定义中“故意”的特征很重要。

问题8-24：病毒的定义是什么？

为了理解病毒的定义，需要注意以下几个问题。

1. 我国政府相关法规中对病毒的定义

我国1994年正式颁布的《中华人民共和国计算机信息系统安全保护条例》对计算机病毒的定义是：编制或插入的破坏计算机功能或者毁坏数据，影响计算机使用，并能自我复制的一组计算机指令或者程序代码。

2. 病毒的主要特征

自我复制传播功能是病毒的主要特征，而病毒的宿主一般是各种可执行的文件。病毒程序专门修改其他宿主文件或硬盘的引导区，来复制自己的恶意程序。在很多情况下，目标文件被修改后并将恶意代码复制进去。一旦感染病毒，宿主文件就变成病毒再去感染其他文件。

问题8-25：病毒、蠕虫、特洛伊木马与流氓软件的主要区别是什么？

为了回答这个问题，可以从以下几个方面分析。

1. 蠕虫的特点

蠕虫是一种复杂的自身复制程序,它完全依靠自身来传播。蠕虫典型的传播方式是利用广泛使用的应用程序,如电子邮件、聊天室等。蠕虫可以将自己附在一封要发送出的邮件上,或者在两个互相信任的系统之间,通过一条简单的 FTP 命令来传播。蠕虫一般不寄生在其他文件或引导区中。

蠕虫和病毒的区别主要表现在以下几个方面。

(1) 蠕虫是独立的程序,而病毒是寄生到其他程序的一段程序。

(2) 蠕虫是通过漏洞进行传播,而病毒是通过复制自身到宿主文件来实现传播。

(3) 蠕虫感染计算机,而病毒感染的是文件系统。

(4) 蠕虫会造成网络拥塞甚至瘫痪,而病毒破坏计算机的文件系统。

(5) 防范蠕虫可以通过及时打补丁、补上漏洞的方法,而防治病毒需要依靠杀毒软件。

2. 木马的特点

木马程序是专为欺骗用户,让用户以为它是友好程序而设计的。木马程序不改变或感染其他的文件,它只是伪装成一种正常程序,随着其他的一些应用程序,装到用户计算机中,但是用户并不知道程序是什么,以欺骗手段诱使用户去激活木马程序。很多木马程序就是后面程序。

蠕虫与木马之间的区别如下。

(1) 木马不对自身进行复制,而蠕虫大量对自身进行复制。

(2) 木马依靠骗取用户的信任去激活它,而蠕虫从一个系统传播到另一个系统,而不需要用户的介入。

3. 流氓软件的特点

流氓软件是指:在未明确提示用户或未经用户许可的情况下,在用户计算机或数字终端设备上安装运行,侵犯用户合法权益的恶意代码软件。

流氓软件的特点主要表现在以下几个方面。

1) 强制安装,难以卸载

在未明确提示用户或未经用户许可的情况下,在用户计算机或其他终端上安装软件;未提供通用的卸载方式,或在不受其他软件影响、人为破坏的情况下,卸载后程序仍然活动。

2) 浏览器劫持

修改用户浏览器或其他相关设置,迫使用户访问特定网站或导致用户无法正常上网。

3) 广告弹出

在未明确提示用户或未经用户许可的情况下,利用安装在用户计算机的软件弹出广告。

4) 恶意收集用户信息

未明确提示用户或未经用户许可的情况下,恶意收集用户信息。

5) 恶意卸载

在未明确提示用户、未经用户许可的情况下,误导、欺骗用户卸载非恶意软件。

流氓软件主要包括恶意广告软件、间谍软件、恶意共享软件。这种软件采用多种手段强行安装和对抗删除。很多用户投诉是在不知情的情况下遭到安装,而其多种反卸载和自动

恢复技术给网络用户造成很大的困扰。它与病毒、木马、蠕虫、垃圾邮件都是有区别的。

问题8-26:垃圾邮件的定义是什么?

为了理解垃圾邮件的定义,需要注意以下几个问题。

1994年,美国的Cantor与Siegel同时在6000个新闻组中发布"绿色抽奖"的广告邮件的行为引发了垃圾邮件的灾难。目前,网络垃圾邮件几乎达到失控的地步,成为互联网中仅次于病毒的第二大公害。

1. 垃圾邮件的特征

垃圾邮件的特征主要表现为未经授权、数量巨大与商业目的三个方面。

2. 垃圾邮件的定义

《中国互联网协会反垃圾邮件规范》中对垃圾邮件的定义如下。

(1) 收件人事先没有提出要求或者不同意接收的广告、电子刊物、各种形式的宣传品等宣传性的电子邮件。

(2) 收件人无法拒收的电子邮件。

(3) 隐藏发件人身份、地址、标题等信息的电子邮件。

(4) 含有虚假的信息源、发件人、路由等信息的电子邮件。

3. 垃圾邮件的危害

由于电子邮件成本低廉,因此许多公司和组织将电子邮件作为其主要的营销与广告宣传手段,它直接导致了垃圾邮件的泛滥。垃圾邮件泛滥的危害主要表现在以下几个方面。

(1) 垃圾邮件的仓储、传输占有了大量的网络存储资源与传输带宽,降低了网络运行效率,影响了正常的电子邮件服务。

(2) 垃圾邮件严重地干扰了用户的正常生活,侵占了收件人大量的邮箱空间。垃圾邮件的清理需要耗费用户大量的时间、精力与费用,大大地降低了电子邮件使用效率与信任度。

(3) 垃圾邮件经常包含大量诈骗、色情,甚至是反动的内容,对社会形成了危害。

(4) 垃圾邮件经常包含病毒、网络钓鱼程序等,成为病毒传播的重要载体,严重地威胁着互联网的安全。

(5) 垃圾邮件常常被攻击者利用,用来造成某些电子邮件服务器的瘫痪与资源耗尽,严重地影响着电子邮件系统的正常运行。

(6) 垃圾邮件的泛滥严重地影响着ISP的服务质量与形象,甚至使它们的IP地址因大量转发垃圾邮件而被封杀。

因此,如何有效解决垃圾邮件所造成的危害,已经成为网络安全研究的一个重要课题。

问题8-27:网络防病毒软件应用的基本方法是什么?

网络防病毒可以从以下两方面入手:一是工作站;二是服务器。为了防止病毒从工作站侵入,可以采取以下措施:使用无盘工作站、带防病毒芯片的网卡、单机防病毒卡或网络防病毒软件。目前,用于网络的防病毒软件有很多,其中多数运行在文件服务器上,可以同时检查服务器和工作站病毒。由于实际局域网中可能有多个服务器,为了方便多服务器的网络管理工作,可将多个服务器组织在一个"域"中,网络管理员只需在域中的

主服务器上设置扫描方式与扫描选项，就可以检查域中多个服务器或工作站是否带病毒。

网络防病毒软件的基本功能是：对服务器和工作站进行扫描、检查、隔离与报警，当发现病毒时，由网络管理员负责清除病毒。网络防病毒软件一般允许用户设置三种扫描方式：实时扫描、预置扫描与人工扫描。实时扫描方式要求连续不断地扫描从文件服务器读/写的文件是否带毒；预置扫描方式可以在预先选择的日期和时间扫描服务器，预置的扫描频度可以是每天一次、每周一次或每月一次，时间最好选择在网络工作不繁忙的时候；人工扫描方式可以在任何时候要求扫描指定的卷、目录和文件。

当网络防病毒软件在服务器上发现病毒后，扫描结果可以保存在查毒记录文件中，并通过两种方法处理染毒文件。一种方法是更改染毒文件的扩展名，使用户无法找到染毒文件，同时提示网络管理员对染毒文件进行消毒，然后将消毒后的文件移回到原目录下；另一种方法是将染毒文件移到特殊的目录下，然后对染毒文件进行消毒处理。一个完整的网络防病毒系统通常由以下几个部分组成：客户端防毒软件、服务器端防毒软件、针对群件的防毒软件、针对黑客的防毒软件。其中，客户端防毒软件除了可以检查一般文件外，还可以检查用ZIP、ARJ 等压缩软件压缩的文件；服务器端防毒软件主要作用是保护服务器，并防止病毒在用户局域网内部传播；针对黑客的防毒软件可通过 MAC 地址与权限列表中的严格匹配，控制可能出现的用户超越权限的行为。

第三部分　习题参考答案

1. 易位密码法加密后的密文为：yknortwcicndounwse
2. ① 私钥
   ② 公钥
   ③ 私钥(加密)
   ④ 公钥(解密)
3. (1) ① 201.1.2.3　② 212.10.5.2　③ 20.16.25.30
   ④ 190.2.2.2　⑤ 201.1.2.3　⑥ 212.10.5.2
   (2) IP 分组头，高层数据
   (3) ESP 头，IP 分组头，高层数据
4. ① K0
   ② K2
   ③ K0
   ④ K1
   ⑤ K0
   ⑥ K0
5. S-B1-S-B1 防火墙层次结构示意图如图 8-19 所示。

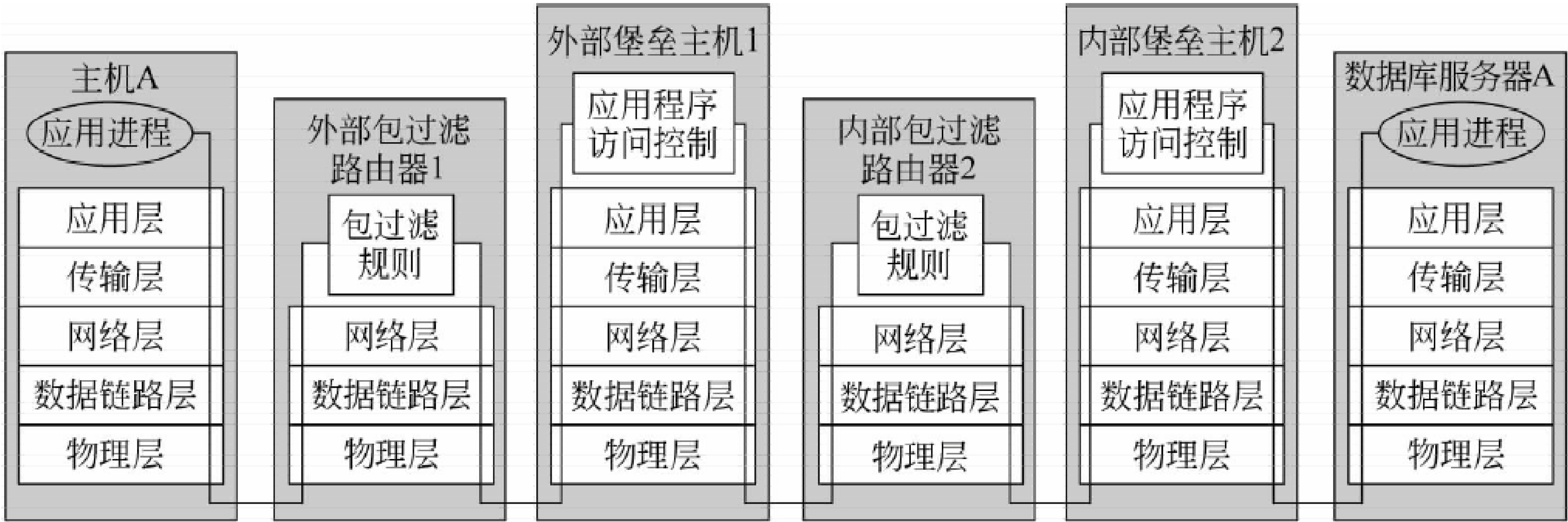


图 8-19　S-B1-S-B1 防火墙层次结构示意图

6. 符合对 ICMP 报文过滤规则要求的防火墙规则表如表 8-2 所示。

表 8-2　符合对 ICMP 报文过滤规则要求的防火墙规则

规则	传输方向	传输协议	源 IP 地址	目的 IP 地址	ICMP 报文类型	动作
1	进入	ICMP	*	*	源主机抑制	允许
2	输出	ICMP	172.16.1.2/24	*	回应请求	允许
3	进入	ICMP	*	172.16.1.2/24	回应应答	允许
4	进入	ICMP	*	172.16.1.2/24	目的主机不可达	允许
5	进入	ICMP	*	172.16.1.2/24	协议不可达	允许
6	进入	ICMP	*	172.16.1.2/24	超时	允许
7	进入	ICMP	*	172.16.1.2/24	回应请求	阻断
8	进入	ICMP	*	172.16.1.2/24	重定向	阻断
9	输出	ICMP	172.16.1.2/24	*	回应请求	阻断
10	输出	ICMP	172.16.1.2/24	*	TTL 超时	阻断

参 考 文 献

[1] Andrew S Tanenbaum. Computer Networks(5th Edition). Prentice-Hall PTR,2011.

[2] Behrouz A Forouzan,Sophia Chung Fegan. TCP/IP Protocol Suite. McGraw-Hill Inc. ,2000.

[3] Eric A Hall. Internet Core Protocols: The Definitive Guide. O'Reilly & Associates Inc. ,2000.

[4] 吴功宜,等.计算机网络高级教程(第 2 版).北京:清华大学出版社,2015.

[5] 吴功宜,等.计算机网络高级软件编程技术(第 2 版).北京:清华大学出版社,2011.

[6] 吴功宜,等.网络安全高级软件编程技术.北京:清华大学出版社,2007.

[7] 吴功宜.计算机网络与互联网技术研究、应用和产业发展.北京:清华大学出版社,2008.

[8] 吴功宜,等.计算机网络技术教程 自顶向下分析与设计方法.北京:机械工业出版社,2009.

[9] 吴功宜,等.计算机网络课程设计(第 2 版).北京:机械工业出版社,2015.

[10] [美]W. Richard Stevens 著,吴英等译. TCP/IP 详解 卷 1:协议.北京:机械工业出版社,2016.

[11] Matthew S Gast. 802.11 无线网络权威指南.南京:东南大学出版社,2007.

[12] [美]Charles E. Spurgeon. Joann Zimmerman 著,蔡仁君译.以太网权威指南.北京:人民邮电出版社,2016.